PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:	WO 93/12141
C07K 13/00, C12P 21/06	A1	(43) International Publication Date:	24 June 1993 (24.06.93)

(21) International Application Number: PCT/US91/09240

(22) International Filing Date: 11 December 1991 (11.12.91)

(71) Applicant: YALE UNIVERSITY [US/US]; 216 Prospect Street, New Haven, CT 06511 (US).

(72) Inventors: FLEMING, Robert, J. ; 70 Coral Way, Rochester, NY 14618 (US). ARTAVANIS-TSAKONAS, Spyridon ; 192 Ridgewood Avenue, Hamden, CT 06517 (US).

(74) Agent: MISROCK, S., Leslie; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated State: CA.

Published
With international search report.

(54) Title: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE SERRATE GENE AND METHODS BASED THEREON

(57) Abstract

The present invention relates to nucleotide sequences of *Serrate* genes, and amino acid sequences of their encoded proteins, as well as derivatives (e.g., fragments) and analogs thereof. In a specific embodiment, the Serrate protein is a human protein. The invention relates to Serrate derivatives and analogs of the invention which are functionally active, i.e., they are capable of displaying one or more known functional activities associated with the full-length (wild-type) Serrate protein. The invention further relates to fragments (and derivatives and analogs thereof) of Serrate which comprise one or more domains of the Serrate protein, including but not limited to the intracellular domain, extracellular domain, transmembrane region, membrane-associated region, or one or more EGF-like repeats of a Serrate protein, or any combination of the foregoing. Antibodies to Serrate, its derivatives and analogs, are additionally provided. Methods of production of the Serrate proteins, derivatives and analogs, e.g., by recombinant means, are also provided.

NUCLEOTIDE AND PROTEIN SEQUENCES OF THE SERRATE GENE AND METHODS BASED THEREON

This invention was made in part with government support under Grant numbers GM 29093 and NS 26084 awarded by the Department of Health and Human Services. The government has certain rights in the invention.

1. INTRODUCTION

The present invention relates to Serrate genes and their encoded protein products. The invention also relates to derivatives and analogs of the Serrate protein. Production of Serrate proteins, derivatives, and antibodies is also provided.

2. BACKGROUND OF THE INVENTION

Genetic analyses in Drosophila have been extremely useful in dissecting the complexity of developmental pathways and identifying interacting loci. However, understanding the precise nature of the processes that underlie genetic interactions requires a knowledge of the protein products of the genes in question.

Recent embryological, genetic and molecular evidence indicates that the early steps of ectodermal differentiation in Drosophila depend on cell interactions (Doe and Goodman, 1985, Dev. Biol. 111:206-219; Technau and Campos-Ortega, 1986, Dev. Biol. 195:445-454; Vässin et al., 1985, J. Neurogenet. 2:291-308; de la Concha et al., 1988, Genetics 118:499-508; Xu et al., 1990, Genes Dev. 4:464-475; Artavanis-Tsakonas, 1988, Trends Genet. 4:95-100). Mutational analyses reveal a small group of zygotically-acting genes, the so called neurogenic loci, which affect the choice of ectodermal cells

between epidermal and neural pathways (Poulson, 1937, Proc. Natl. Acad. Sci. 23:133-137; Lehmann et al., 1983, Wilhelm Roux's Arch. Dev. Biol. 192:62-74; Jürgens et al., 1984, Wilhelm Roux's Arch. Dev. Biol. 193:283-295; Wieschaus et al., 1984, Wilhelm Roux's Arch. Dev. Biol. 193:296-307; Nüsslein-Volhard et al., 1984, Wilhelm Roux's Arch. Dev. Biol. 193:267-282). Null mutations in any one of the zygotic neurogenic loci -- Notch (N), Delta (Dl), mastermind (mam), Enhancer of Split (E(spl), neuralized (neu), and big brain (bib) -- result in hypertrophy of the nervous system at the expense of ventral and lateral epidermal structures. This effect is due to the misrouting of epidermal precursor cells into a neuronal pathway, and implies that neurogenic gene function is necessary to divert cells within the neurogenic region from a neuronal fate to an epithelial fate. Serrate has been identified as a genetic unit capable of interacting with the Notch locus (Xu et al., 1990, Genes Dev. 4:464-475). These genetic and developmental observations have led to the hypothesis that the protein products of the neurogenic loci function as components of a cellular interaction mechanism necessary for proper epidermal development (Artavanis-Tsakonas, S., 1988, Trends Genet. 4:95-100).

Mutational analyses also reveal that the action of the neurogenic genes is pleiotropic and is not limited solely to embryogenesis. For example, ommatidial, bristle and wing formation, which are known also to depend upon cell interactions, are affected by neurogenic mutations (Morgan et al., 1925, Bibliogr. Genet. 2:1-226; Welshons, 1956, Dros. Inf. Serv. 30:157-158; Preiss et al., 1988, EMBO J. 7:3917-3927; Shellenbarger and Mohler, 1978, Dev. Biol. 62:432-446; Technau and Campos-Ortega, 1986, Wilhelm

Roux's Dev. Biol. 195:445-454; Tomlison and Ready, 1987, Dev. Biol. 120:366-376; Cagan and Ready, 1989, Genes Dev. 3:1099-1112).

Sequence analyses (Wharton et al., 1985, Cell 43:567-581; Kidd and Young, 1986, Mol. Cell. Biol. 6:3094-3108; Vässin, et al., 1987, EMBO J. 6:3431-3440; Kopczynski, et al., 1988, Genes Dev. 2:1723-1735) have shown that two of the neurogenic loci, Notch and Delta, appear to encode transmembrane proteins that span the membrane a single time. The Notch gene encodes a ~300 kd protein (we use "Notch" to denote this protein) with a large N-terminal extracellular domain that includes 36 epidermal growth factor (EGF)-like tandem repeats followed by three other cysteine-rich repeats, designated Notch/lin-12 repeats (Wharton, et al., 1985, Cell 43:567-581; Kidd and Young, 1986, Mol. Cell. Biol. 6:3094-3108; Yochem, et al., 1988, Nature 335:547-550). Delta encodes a ~100 kd protein (we use "Delta" to denote DLZM, the protein product of the predominant zygotic and maternal transcripts; Kopczynski, et al., 1988, Genes Dev. 2:1723-1735) that has nine EGF-like repeats within its extracellular domain (Vässin, et al., 1987, EMBO J. 6:3431-3440; Kopczynski, et al., 1988, Genes Dev. 2:1723-1735). Molecular studies have lead to the suggestion that Notch and Delta constitute biochemically interacting elements of a cell communication mechanism involved in early developmental decisions (Fehon et al., 1990, Cell 61:523-534).

The EGF-like motif has been found in a variety of proteins, including those involved in the blood clotting cascade (Furie and Furie, 1988, Cell 53: 505-518). In particular, this motif has been found in extracellular proteins such as the blood

clotting factors IX and X (Rees et al., 1988, EMBO J. 7:2053-2061; Furie and Furie, 1988, Cell 53: 505-518), in other Drosophila genes (Knust et al., 1987 EMBO J. 761-766; Rothberg et al., 1988, Cell 55:1047-1059), and in some cell-surface receptor proteins, such as thrombomodulin (Suzuki et al., 1987, EMBO J. 6:1891-1897) and LDL receptor (Sudhof et al., 1985, Science 228:815-822). A protein binding site has been mapped to the EGF repeat domain in thrombomodulin and urokinase (Kurosawa et al., 1988, J. Biol. Chem 263:5993-5996; Appella et al., 1987, J. Biol. Chem. 262:4437-4440).

Citation of references hereinabove shall not be construed as an admission that such references are prior art to the present invention.

3. SUMMARY OF THE INVENTION

The present invention relates to nucleotide sequences of Serrate genes, and amino acid sequences of their encoded proteins, as well as derivatives (e.g., fragments) and analogs thereof. Nucleic acids hybridizable to or complementary to the foregoing nucleotide sequences are also provided. In a specific embodiment, the Serrate protein is a human protein.

The invention relates to Serrate derivatives and analogs of the invention which are functionally active, i.e., they are capable of displaying one or more known functional activities associated with a full-length (wild-type) Serrate protein. Such functional activities include but are not limited to antigenicity [ability to bind (or compete with Serrate for binding) to an anti-Serrate antibody], immunogenicity (ability to generate antibody which binds to Serrate), ability to bind (or compete with Serrate for binding) to Notch or other toporythmic

proteins or fragments thereof ("adhesiveness"), ability to bind (or compete with Serrate for binding) to a receptor for Serrate. "Toporythmic proteins" as used herein, refers to the protein products of Notch, Delta, Serrate, Enhancer of split, and Deltex, as well as other members of this interacting gene family which may be identified, e.g., by virtue of the ability of their gene sequences to hybridize, or their homology to Delta, Serrate, or Notch, or the ability of their genes to display phenotypic interactions.

The invention further relates to fragments (and derivatives and analogs thereof) of Serrate which comprise one or more domains of the Serrate protein, including but not limited to the intracellular domain, extracellular domain, transmembrane domain, membrane-associated region, or one or more EGF-like (homologous) repeats of a Serrate protein, or any combination of the foregoing.

Antibodies to Serrate, its derivatives and analogs, are additionally provided.

Methods of production of the Serrate proteins, derivatives and analogs, e.g., by recombinant means, are also provided.

3.1. DEFINITIONS

As used herein, underscoring the name of a gene shall indicate the gene, in contrast to its encoded protein product which is indicated by the name of the gene in the absence of any underscoring. For example, "Serrate" shall mean the Serrate gene, whereas "Serrate" shall indicate the protein product of the Serrate gene.

4. DESCRIPTION OF THE FIGURES

Figure 1. Phenotypic interactions between Notch and Serrate. (a) w^aspl wing blade showing characteristic wild-type symmetry, venation, and marginal wing bristles and hairs. (b) nd/Y male. Distal wing notches and loss of posterior hairs are evident. (c) Ser^D/+ heterozygote. Note similarity to nd/Y wing blade in Fig. 1b. (d) nd/Y; Ser^D/+ transheterozygote wing blade. Mutant wing shows typical "fig leaf" shape, distorted wing veins, and loss of the majority of marginal bristles and hairs, with the exception of the anterodistal wing margin. (e) +/Y; Ser^D/Dp(3R)CosP479BE (N^+) male. The extra N^+ copy suppresses the heterozygous Ser^D dominant phenotype (compare to Fig. 1c). Also note suppression of the Confluens phenotype (see text). (f) Ser^D/Ser^D homozygote. Note the increased severity of the phenotype relative to Ser^D/+ (compare to Fig. 1c).

Figure 2. Phenotypes of Serrate lethal mutations. (a-d) Cuticular preparations; (e-h) anti-HRP preparations. (a) Wild-type cuticular pattern of embryo just prior to hatching. (b) Typical non-retracted germ band $Ser^{revertant}$ homozygote. Note position of Filzkörper (arrow). (c) $Ser^{revertant}$ homozygote lacking the cuticle of the cephalic regions and the first and second thoracic segments. (d) Severely affected $Ser^{revertant}$ homozygote displaying limited cuticle differentiation. (e) Wild-type embryo showing typical nervous system differentiation. (f) Homozygous $Ser^{revertant}$ embryo displaying a "mild" disruption of the longitudinal and commissural axon tracts. (g) $Ser^{revertant}$ homozygote; note singular, twisted longitudinal connective. (h) Severely affected $Ser^{revertant}$ homozygote with dispersed clumps of neural material remaining.

Figure 3. Molecular map of the Serrate-encoding region. Approximately 85 kb of cloned

genomic DNA from the 97F chromosomal region are presented along with the restriction sites of three enzymes [(B) BamHI; (E) EcoRI; (H) HindIII]. The locations of individual DNA alterations associated with Serrate allelic breakpoints are displayed above the genomic DNA (for descriptions of mutant alleles, see Section 6, infra; (rev 3 and rev 2-11) $Ser^{rev\ 3}$ and $Ser^{rev\ 2\text{-}11}$, respectively; (R128) T(Y:3)R128. The shaded box from coordinates 0 to +3 represents the region of EGF homology detectable by Southern hybridization. The BamHI site adjacent to the EGF homology was arbitrarily chosen as position 0. Map orientation is with the centromere to the left. At the bottom of the figure are shown the individual recombinant phage isolates. The C1 and C3 cDNAs together constitute the larger of the two Serrate messages (~5.6 kb). Intron positions and coding capacities have been approximated solely upon cross hybridization of the cDNAs with the genomic DNA regions.

Figure 4. Serrate sequence analysis. The complete 5561 bp sequence (SEQ ID NO:1) derived from cDNAs C1 and C3 is shown. Nucleotide numbering is at left, amino acid numbering of the predicted open reading frame (ORF) is at right. The deduced protein product appears to be a transmembrane protein of 1404 amino acids (SEQ ID NO:2). Hydrophobic regions are denoted inside dashed boxes; amino acids 51 to 80 represent the likely signal peptide; amino acids 542 to 564 represent the potential membrane associated region; amino acids 1221 to 1245 represent the putative transmembrane domain. The first cysteine of each of the fourteen EGF-like repeats is denoted with a solid black box, and each repeat is underlined. The partial EGF-like repeat is considered "degenerate," since the fourth cysteine residue of this repeat has

been changed to lysine (shown in boldface type at amino acid position 268). The initial cysteine of this repeat is denoted with an open box (amino acid 284), and the repeat is underlined. Amino acid insertions occur in the fourth, sixth, and tenth EGF-like repeats and are denoted by hatched underlines.

Figure 5. The Serrate transcript and deduced protein product. (a) The composite transcript shown was constructed from the C1 and C3 cDNAs, which overlap by 109 bp. Selected restriction enzyme cleavage sites are shown. The hatched box represents the 4212 bp ORF. Open boxes represent the 442 bp 5'-untranslated leader and 900 bp 3'-trailer sequence. (b) Kyte-Doolittle hydropathy plot of the deduced 1404 amino acid protein. (SP) Putative signal peptide; (MA) potential membrane associated region; (TM) likely transmembrane domain. (c) Cartoon representation of the gross structural features of the predicted Serrate protein. The darkly shaded region, including the partial EGF-like repeat (PR) is ~250 amino acids in length and homologous to the Delta protein. Bracketed EGF-like repeats labelled (A, B, and C) contain insertions of amino acids and thus differ from the canonical EGF-like structure. Other features of the protein include the signal peptide (SP), a cysteine rich region, a transmembrane domain (TM), and an intracellular region of ~160 amino acids.

Figure 6. Temporal profile of Serrate transcript accumulation. Each lane contains five μg of poly(A)$^+$ RNA. The stage of the embryonic RNAs is denoted in hours after egg laying; (L1, L2, and L3) RNA from the first, second and third larval instar periods; (EP and LP) early and late pupal stages; (M and F) adult male and female RNAs, respectively. A composite cDNA subclone (constructed from C1 and C3)

was used as a hybridization probe. Serrate transcription is represented primarily as a 5.5 kb and 5.6 kb doublet beginning at 4-8 hours of embryogenesis. A transient 3.4 kb transcript is observed only during 2-4 hr of embryogenesis. The pupal and adult RNAs were fractionated on a separate gel for a longer period of time for better resolution. Equivalent loadings of RNA were noted by ethidium bromide staining of the RNA gels and confirmed by subsequent probing with an actin 5C probe shown at bottom; Fyrberg et al., 1983, Cell 33:115-123). Minor bands were not consistently observed in other blots and may reflect other EGF-homologous transcripts

Figure 7. Whole-mount in situ Serrate transcripts. Embryos are oriented with anterior to the left and dorsal side up unless otherwise noted. (a) Dorsal view of an early stage 10 embryo (mid-dorsal focal plane). Earliest expression occurs in the ectoderm of the foregut (FG) and presumptive clypeolabrum (CL). (b) Dorsal view of a germ band-extended embryo (late stage 10). Additional expression occurs near the proctodeum (PR), within the eighth (A8) and ninth (A9) abdominal segments, and in the labial and maxillary primordia (arrow). (c) Lateral view of an early stage 11 embryo. The lateral (LE) and ventral (VE) expression patterns are out of register and do not include the tracheal pits (TP). (d) Germ band-extended embryo (mid stage 11) dissected and flattened such that the dorsal surfaces are at the lateral edges. Extensive expression is observed between the labial (LB), maxillary (MX), and mandibular (MN) lobes, and within the hypopharynx (HP) and clypeolabrum (CL). Expression is also apparent in the salivary gland placodes (SP) that have moved to the ventral midline. Note relationship between

lateral and ventral patterns and elaboration of expression in the tail region [presumptive telson (TL)]. (e) Germ band-retracting embryo (stage 12; lateral view). Lateral expression (LE) is beginning to coalesce. (f) Lateral view of a germ band-retracted embryo (stage 13). The lateral expression is beginning to extend both dorsally and ventrally in each thoracic and abdominal segment and is most pronounced in the first thoracic segment (T1). A portion of the lateral expression now appears to include the presumptive trachea (T). Ventrally, note different expression (VE) patterns in the thoracic versus abdominal segments. (g) Lateral view of an early stage 14 embryo. Outline of the presumptive trachea (T) is distinct from the overlying epidermal expression. Arrows denote the zigzag pattern of lateral expression. (h) Dissected embryo (stage 14) opened along the dorsal midline and laid flat. Two areas of hindgut expression (HG1 and HG2) are apparent; HG1 occurs near the origin of the Malpighian tubules. (i) Ventral view of a stage-16 embryo focusing on the ventral nerve cord (VNC). Earlier expression in the salivary gland placodes (SP in panel d) now constitutes the SD. Expression in the proventriculus (PV) and the maxillary/mandibular region (MX/MN) is slightly out of focus. (j) Dorsomedial focal plane of same embryo as in (i); head involution is nearly complete. The in-pocketings of expression in the thoracic segments (T1, T2, and T3) may represent imaginal disc primordia. Pharyngeal expression (PH) is a combination of clypeolabrum and hypopharyngeal expression noted earlier. (k) Dorsal view of the same embryo as in (i) and (j). Note individual expressing cells in the brain lobes (BC). Expression in the fully differentiated trachea (T) and

hindgut (H1) is evident. (1) Flattened preparation of early stage 16 embryo. Expression within the telson (TL) now constitutes a ring around the presumptive anal pads.

Figure 8. Amino acid comparison of amino-terminal Serrate-Delta homology. Conserved regions are indicated at the top of the figure (* = identical amino acids; ' = conservative changes in sequence). Serrate (see SEQ ID NO:2) is shown above line, Delta (SEQ ID NO:4) below. The sequence begins at Serrate amino acid position 59; the partial EGF-like repeat of both Serrate and Delta is boxed. The Serrate amino acid sequence (amino acids 79-282 of Fig. 4) placed into the chimeric ΔEGF Notch construct and determined to be sufficient for Notch binding is presented in boldface type. The positions of the synthetic degenerate primers (designated FLE1 through FLE4R) are shown; refer to Figure 9 for nucleotide composition.

Figure 9. Nucleotide comparison of amino-terminal Serrate-Delta homology. The nucleotide sequence corresponding to the amino acid sequence in Figure 8 is shown (Serrate sequence: see SEQ ID NO:1; Delta sequence: SEQ ID NO:3). The DNA encoding the partial EGF-repeat is boxed. The Serrate nucleotide sequence (nucleotides 676-1287 of Fig. 4) placed into the chimeric ΔEGF Notch construct determined to be sufficient for Notch binding is presented in boldface type. The sequences of the synthetic degenerate primers (designated FLE1 through FLE4R) are given against black backgrounds.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to nucleotide sequences of Serrate genes, and amino acid sequences of their encoded proteins. The invention further

relates to fragments and other derivatives, and analogs, of Serrate proteins. Nucleic acids encoding such fragments or derivatives are also within the scope of the invention. In a preferred embodiment of the invention, the Serrate protein is a human protein. Production of the foregoing proteins and derivatives, e.g., by recombinant methods, is provided.

The invention relates to Serrate derivatives and analogs of the invention which are functionally active, i.e., they are capable of displaying one or more known functional activities associated with a full-length (wild-type) Serrate protein. Such functional activities include but are not limited to antigenicity [ability to bind (or compete with Serrate for binding) to an anti-Serrate antibody], immunogenicity (ability to generate antibody which binds to Serrate), ability to bind (or compete with Serrate for binding) to Notch or other toporythmic proteins or fragments thereof ("adhesiveness"), ability to bind (or compete with Serrate for binding) to a receptor for Serrate. "Toporythmic proteins" as used herein, refers to the protein products of Notch, Delta, Serrate, Enhancer of split, and Deltex, as well as other members of this interacting gene family which may be identified, e.g., by virtue of the ability of their gene sequences to hybridize, or their homology to Delta, Serrate, or Notch, or the ability of their genes to display phenotypic interactions.

The invention further relates to fragments (and derivatives and analogs thereof) of Serrate which comprise one or more domains of the Serrate protein, including but not limited to the intracellular domain, extracellular domain, transmembrane domain, membrane-associated region, or one or more EGF-like

(homologous) repeats of a Serrate protein, or any combination of the foregoing.

Antibodies to Serrate, its derivatives and analogs, are additionally provided.

As demonstrated infra (see Section 6), Serrate plays a critical role in development and other physiological processes. The nucleic acid and amino acid sequences and antibodies thereto of the invention can be used for the detection and quantitation of Serrate mRNA of human and other species, to study expression thereof, to produce Serrate and fragments and other derivatives and analogs thereof, in the study and manipulation of differentiation and other physiological processes, and may be of therapeutic or diagnostic use.

The invention is illustrated by way of examples infra which disclose, inter alia, the cloning and sequencing of D. melanogaster Serrate (Section 6); the construction and recombinant expression of a Serrate chimeric/fusion derivative and production of antibodies thereto (Section 7); the recombinant expression of Serrate, a Serrate fragment lacking the EGF-like repeats present in Serrate, and a chimeric Notch-Serrate derivative, and assays for binding to Notch (Section 8); and the cloning of a human Serrate homolog (Section 9).

For clarity of disclosure, and not by way of limitation, the detailed description of the invention will be divided into the following sub-sections:

(i) Isolation of the Serrate Gene;

(ii) Expression of the Serrate Gene;

(iii) Identification and Purification of the Expressed Gene Product;

(iv) Structure of the Serrate Gene and Protein;

(v) Generation of Antibodies to Serrate Proteins and Derivatives Thereof;

(vi) Serrate Derivatives and Analogs;

(vii) Assays of Serrate Proteins, Derivatives, and Analogs.

5.1. ISOLATION OF THE SERRATE GENE

The invention relates to the nucleotide sequences of Serrate consisting of at least 8 nucleotides (i.e., a hybridizable portion). In a specific embodiment, the invention relates to the nucleic acid sequence of the human Serrate gene. In another embodiment, the invention relates to the Drosophila Serrate gene. In a preferred, but not limiting, aspect of the invention, a Drosophila Serrate DNA sequence (ATCC Accession Number _____) can be cloned and sequenced by the method described in Section 6, infra. The invention also relates to nucleic acids hybridizable to or complementary to the foregoing sequences.

Nucleic acids encoding fragments and derivatives of Serrate (see Section 5.6) are additionally provided.

Fragments of Serrate nucleic acids comprising regions of homology to other toporythmic proteins are also provided. For example, the total region of homology with Delta spans nucleotides 627-1290 (Fig. 4) (see SEQ ID NO:1) of the Serrate sequence. Nucleic acids encoding conserved regions between Delta and Serrate, such as those represented by Serrate amino acids 63-73, 124-134, 149-158, 195-206, 214-219, and 250-259 (see SEQ ID NO:2), are also provided.

A preferred embodiment for the cloning of human Serrate, presented as a particular example but not by way of limitation, follows:

A human expression library is constructed by methods known in the art. For example, human mRNA is isolated, cDNA is made and ligated into an expression vector (e.g., a bacteriophage derivative) such that it is capable of being expressed by the host cell into which it is then introduced. Various screening assays can then be used to select for the expressed human Serrate product. In one embodiment, anti-Serrate antibodies can be used for selection.

In another preferred aspect, PCR is used to amplify the desired sequence in the library, prior to selection. Oligonucleotide primers representing known Serrate sequences can be used as primers in PCR. In a preferred aspect, the oligonucleotide primers represent at least part of the Serrate conserved segments of strong homology between Serrate and Delta. For example, oligonucleotides may be obtained corresponding to parts of the four highly conserved regions between Delta and Serrate, i.e., that represented by Serrate AA 124-134, 149-158, 214-219, and 250-259 (see SEQ ID NO:2). The synthetic oligonucleotides may be utilized as primers to amplify by PCR sequences from a source (RNA or DNA), preferably a cDNA library, of potential interest. In a specific embodiment an oligonucleotide primer pair used for PCR with a cDNA library is FLE1/FE3R, FLE1/FLE4R, FLE2/FLE3R, or FLE2/FLE4R shown in Fig. 9. (see SEQ ID NO:1) (PCR can be carried out, e.g., by use of a Perkin-Elmer Cetus thermal cycler and Taq polymerase (Gene Amp™)). The DNA being amplified can include mRNA or cDNA or genomic DNA from any eukaryotic species. One can choose to synthesize

several different degenerate primers, for use in the PCR reactions. It is also possible to vary the stringency of hybridization conditions used in priming the PCR reactions, to allow for greater or lesser degrees of nucleotide sequence similarity between the Serrate homolog and the known Serrate. After successful amplification of a segment of a Serrate homolog, that segment may be molecularly cloned and sequenced, and utilized as a probe to isolate a complete cDNA or genomic clone. This, in turn, will permit the determination of the gene's complete nucleotide sequence, the analysis of its expression, and the production of its protein product for functional analysis, as described infra. In this fashion, additional genes encoding Serrate proteins may be identified. Such a procedure is presented by way of example in Section 9, infra.

The above-methods are not meant to limit the following general description of methods by which clones of Serrate may be obtained.

Any eukaryotic cell potentially can serve as the nucleic acid source for the molecular cloning of the Serrate gene. The nucleic acid sequences encoding Serrate can be isolated from human, porcine, bovine, feline, avian, equine, canine, as well as additional primate sources, insects, etc. For example, we have amplified fragments of the appropriate size in Drosophila, mouse, Xenopus, and human, by PCR using cDNA libraries with Drosophila Serrate primers. The DNA may be obtained by standard procedures known in the art from cloned DNA (e.g., a DNA "library"), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell. (See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed.,

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II.) Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will contain only exon sequences. Whatever the source, the gene should be molecularly cloned into a suitable vector for propagation of the gene.

In the molecular cloning of the gene from genomic DNA, DNA fragments are generated, some of which will encode the desired gene. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNAse in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The linear DNA fragments can then be separated according to size by standard techniques, including but not limited to, agarose and polyacrylamide gel electrophoresis and column chromatography.

Once the DNA fragments are generated, identification of the specific DNA fragment containing the desired gene may be accomplished in a number of ways. For example, if an amount of a portion of a Serrate (of any species) gene or its specific RNA, or a fragment thereof, e.g., an extracellular domain (see Section 5.6), is available and can be purified and labeled, the generated DNA fragments may be screened by nucleic acid hybridization to the labeled probe (Benton, W. and Davis, R., 1977, Science 196:180; Grunstein, M. And Hogness, D., 1975, Proc. Natl. Acad. Sci. U.S.A. 72:3961). Those DNA fragments with substantial homology to the probe will hybridize. It is also possible to identify the appropriate fragment

by restriction enzyme digestion(s) and comparison of fragment sizes with those expected according to a known restriction map if such is available. Further selection can be carried out on the basis of the properties of the gene. Alternatively, the presence of the gene may be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones, or DNA clones which hybrid-select the proper mRNAs, can be selected which produce a protein that, e.g., has similar or identical electrophoretic migration, isolectric focusing behavior, proteolytic digestion maps, receptor binding activity, in vitro aggregation activity ("adhesiveness") or antigenic properties as known for Serrate. If an antibody to Serrate is available, the Serrate protein may be identified by binding of labeled antibody to the putatively Serrate synthesizing clones, in an ELISA (enzyme-linked immunosorbent assay)-type procedure.

The Serrate gene can also be identified by mRNA selection by nucleic acid hybridization followed by in vitro translation. In this procedure, fragments are used to isolate complementary mRNAs by hybridization. Such DNA fragments may represent available, purified Serrate DNA of another species (e.g., Drosophila). Immunoprecipitation analysis or functional assays (e.g., aggregation ability in vitro; binding to receptor; see infra) of the in vitro translation products of the isolated products of the isolated mRNAs identifies the mRNA and, therefore, the complementary DNA fragments that contain the desired sequences. In addition, specific mRNAs may be selected by adsorption of polysomes isolated from cells to immobilized antibodies specifically directed against Serrate protein. A radiolabelled Serrate cDNA

can be synthesized using the selected mRNA (from the adsorbed polysomes) as a template. The radiolabelled mRNA or cDNA may then be used as a probe to identify the Serrate DNA fragments from among other genomic DNA fragments.

Alternatives to isolating the Serrate genomic DNA include, but are not limited to, chemically synthesizing the gene sequence itself from a known sequence or making cDNA to the mRNA which encodes the Serrate protein. For example, RNA for cDNA cloning of the Serrate gene can be isolated from cells which express Serrate. Other methods are possible and within the scope of the invention.

The identified and isolated gene can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Such vectors include, but are not limited to, bacteriophages such as lambda derivatives, or plasmids such as PBR322 or pUC plasmid derivatives. The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In an alternative method, the cleaved vector and Serrate gene may be

modified by homopolymeric tailing. Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated.

In an alternative method, the desired gene may be identified and isolated after insertion into a suitable cloning vector in a "shot gun" approach. Enrichment for the desired gene, for example, by size fractionization, can be done before insertion into the cloning vector.

In specific embodiments, transformation of host cells with recombinant DNA molecules that incorporate the isolated Serrate gene, cDNA, or synthesized DNA sequence enables generation of multiple copies of the gene. Thus, the gene may be obtained in large quantities by growing transformants, isolating the recombinant DNA molecules from the transformants and, when necessary, retrieving the inserted gene from the isolated recombinant DNA.

5.2. EXPRESSION OF THE SERRATE GENE

The nucleotide sequence coding for a Serrate protein or a functionally active fragment or other derivative thereof (see Section 5.6), can be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. The necessary transcriptional and translational signals can also be supplied by the native Serrate gene and/or its flanking regions. A variety of host-vector systems may be utilized to express the protein-coding sequence. These include but are not limited to mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.);

insect cell systems infected with virus (e.g., baculovirus); microorganisms such as yeast containing yeast vectors, or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used. In a specific embodiment, the adhesive portion of the Serrate gene is expressed. In other specific embodiments, the human Serrate gene is expressed, or a sequence encoding a functionally active portion of human Serrate. In yet another embodiment, a fragment of Serrate comprising the extracellular domain, or other derivative, or analog of Serrate is expressed.

Any of the methods previously described for the insertion of DNA fragments into a vector may be used to construct expression vectors containing a chimeric gene consisting of appropriate transcriptional/translational control signals and the protein coding sequences. These methods may include in vitro recombinant DNA and synthetic techniques and in vivo recombinants (genetic recombination). Expression of nucleic acid sequence encoding a Serrate protein or peptide fragment may be regulated by a second nucleic acid sequence so that the Serrate protein or peptide is expressed in a host transformed with the recombinant DNA molecule. For example, expression of a Serrate protein may be controlled by any promoter/enhancer element known in the art. Promoters which may be used to control toporythmic gene expression include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus

(Yamamoto, et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42); prokaryotic expression vectors such as the β-lactamase promoter (Villa-Kamaroff, et al., 1978, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731), or the tac promoter (DeBoer, et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:21-25); see also "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242:74-94; plant expression vectors comprising the nopaline synthetase promoter region (Herrera-Estrella et al., Nature 303:209-213) or the cauliflower mosaic virus 35S RNA promoter (Gardner, et al., 1981, Nucl. Acids Res. 9:2871), and the promoter of the photosynthetic enzyme ribulose biphosphate carboxylase (Herrera-Estrella et al., 1984, Nature 310:115-120); promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter, and the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, Cell 38:639-646; Ornitz et al., 1986, Cold Spring Harbor Symp. Quant. Biol. 50:399-409; MacDonald, 1987, Hepatology 7:425-515); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, Nature 315:115-122), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., 1984, Cell 38:647-658; Adames et al., 1985, Nature 318:533-538; Alexander et al., 1987, Mol. Cell. Biol. 7:1436-1444), mouse mammary tumor virus control

region which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, Cell 45:485-495), albumin gene control region which is active in liver (Pinkert et al., 1987, Genes and Devel. 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, Mol. Cell. Biol. 5:1639-1648; Hammer et al., 1987, Science 235:53-58; alpha 1-antitrypsin gene control region which is active in the liver (Kelsey et al., 1987, Genes and Devel. 1:161-171), beta-globin gene control region which is active in myeloid cells (Mogram et al., 1985, Nature 315:338-340; Kollias et al., 1986, Cell 46:89-94; myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, Nature 314:283-286), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378).

Expression vectors containing Serrate gene inserts can be identified by three general approaches: (a) nucleic acid hybridization, (b) presence or absence of "marker" gene functions, and (c) expression of inserted sequences. In the first approach, the presence of a foreign gene inserted in an expression vector can be detected by nucleic acid hybridization using probes comprising sequences that are homologous to an inserted toporythmic gene. In the second approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" gene functions (e.g., thymidine kinase activity, resistance to antibiotics, transformation phenotype, occlusion body formation in baculovirus, etc.) caused by the insertion of foreign

genes in the vector. For example, if the Serrate gene is inserted within the marker gene sequence of the vector, recombinants containing the Serrate insert can be identified by the absence of the marker gene function. In the third approach, recombinant expression vectors can be identified by assaying the foreign gene product expressed by the recombinant. Such assays can be based, for example, on the physical or functional properties of the Serrate gene product in vitro assay systems, e.g., aggregation (binding) with Notch, binding to a receptor, binding with antibody.

Once a particular recombinant DNA molecule is identified and isolated, several methods known in the art may be used to propagate it. Once a suitable host system and growth conditions are established, recombinant expression vectors can be propagated and prepared in quantity. As previously explained, the expression vectors which can be used include, but are not limited to, the following vectors or their derivatives: human or animal viruses such as vaccinia virus or adenovirus; insect viruses such as baculovirus; yeast vectors; bacteriophage vectors (e.g., lambda), and plasmid and cosmid DNA vectors, to name but a few.

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus, expression of the genetically engineered Serrate protein may be controlled. Furthermore, different host cells have characteristic and specific mechanisms for the translational and post-translational processing and modification (e.g.,

glycosylation, cleavage [e.g., of signal sequence]) of proteins. Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed. For example, expression in a bacterial system can be used to produce an unglycosylated core protein product. Expression in yeast will produce a glycosylated product. Expression in mammalian cells can be used to ensure "native" glycosylation of a heterologous mammalian toporythmic protein. Furthermore, different vector/host expression systems may effect processing reactions such as proteolytic cleavages to different extents.

Both cDNA and genomic sequences can be cloned and expressed.

5.3. IDENTIFICATION AND PURIFICATION OF THE EXPRESSED GENE PRODUCT

Once a recombinant which expresses the Serrate gene sequence is identified, the gene product can be analyzed. This is achieved by assays based on the physical or functional properties of the product, including radioactive labelling of the product followed by analysis by gel electrophoresis, immunoassay, etc.

Once the Serrate protein is identified, it may be isolated and purified by standard methods including chromatography (e.g., ion exchange, affinity, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. The functional properties may be evaluated using any suitable assay (see Section 5.7).

Alternatively, once a Serrate protein produced by a recombinant is identified, the amino acid sequence of the protein can be deduced from the

nucleotide sequence of the chimeric gene contained in the recombinant. As a result, the the protein can be synthesized by standard chemical methods known in the art (e.g., see Hunkapiller, M., et al., 1984, Nature 310:105-111).

In a specific embodiment of the present invention, such Serrate proteins, whether produced by recombinant DNA techniques or by chemical synthetic methods, include but are not limited to those containing, as a primary amino acid sequence, all or part of the amino acid sequence substantially as depicted in Figure 4 (SEQ ID NO:2), as well as fragments and other derivatives, and analogs thereof.

5.4. STRUCTURE OF THE SERRATE GENE AND PROTEIN

The structure of the Serrate gene and protein can be analyzed by various methods known in the art.

5.4.1. GENETIC ANALYSIS

The cloned DNA or cDNA corresponding to the Serrate gene can be analyzed by methods including but not limited to Southern hybridization (Southern, E.M., 1975, J. Mol. Biol. 98:503-517), Northern hybridization (see e.g., Freeman et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:4094-4098), restriction endonuclease mapping (Maniatis, T., 1982, Molecular Cloning, A Laboratory, Cold Spring Harbor, New York), and DNA sequence analysis. Polymerase chain reaction (PCR; U.S. Patent Nos. 4,683,202, 4,683,195 and 4,889,818; Gyllenstein et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7652-7656; Ochman et al., 1988, Genetics 120:621-623; Loh et al., 1989, Science 243:217-220) followed by Southern hybridization with a Serrate-specific probe can allow the detection of the

Serrate gene in DNA from various cell types. In one embodiment, Southern hybridization can be used to determine the genetic linkage of Serrate. Northern hybridization analysis can be used to determine the expression of the Serrate gene. Various cell types, at various states of development or activity can be tested for Serrate expression. Examples of such techiques and their results are described in Section 6, infra. The stringency of the hybridization conditions for both Southern and Northern hybridization can be manipulated to ensure detection of nucleic acids with the desired degree of relatedness to the specific Serrate probe used.

Restriction endonuclease mapping can be used to roughly determine the genetic structure of the Serrate gene. In a particular embodiment, cleavage with restriction enzymes can be used to derive the restriction map shown in Figure 3, infra. Restriction maps derived by restriction endonuclease cleavage can be confirmed by DNA sequence analysis.

DNA sequence analysis can be performed by any techniques known in the art, including but not limited to the method of Maxam and Gilbert (1980, Meth. Enzymol. 65:499-560), the Sanger dideoxy method (Sanger, F., et al., 1977, Proc. Natl. Acad. Sci. U.S.A. 74:5463), the use of T7 DNA polymerase (Tabor and Richardson, U.S. Patent No. 4,795,699), or use of an automated DNA sequenator (e.g., Applied Biosystems, Foster City, CA). The cDNA sequence of a representative Serrate gene comprises the sequence substantially as depicted in Figure 4, and described in Section 6, infra.

5.4.2. PROTEIN ANALYSIS

The amino acid sequence of the Serrate protein can be derived by deduction from the DNA sequence, or alternatively, by direct sequencing of the protein, e.g., with an automated amino acid sequencer. The amino acid sequence of a representative Serrate protein comprises the sequence substantially as depicted in Figure 4, and detailed in Section 6, infra, with the representative mature protein that shown by amino acid numbers 81-1404.

The Serrate protein sequence can be further characterized by a hydrophilicity analysis (Hopp, T. and Woods, K., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:3824). A hydrophilicity profile can be used to identify the hydrophobic and hydrophilic regions of the Serrate protein and the corresponding regions of the gene sequence which encode such regions. A hydrophilicity profile of the Serrate protein described in the examples section infra is depicted in Figure 5.

Secondary, structural analysis (Chou, P. and Fasman, G., 1974, Biochemistry 13:222) can also be done, to identify regions of Serrate that assume specific secondary structures.

Manipulation, translation, and secondary structure prediction, as well as open reading frame prediction and plotting, can also be accomplished using computer software programs available in the art.

Other methods of structural analysis can also be employed. These include but are not limited to X-ray crystallography (Engstom, A., 1974, Biochem. Exp. Biol. 11:7-13) and computer modeling (Fletterick, R. and Zoller, M. (eds.), 1986, Computer Graphics and Molecular Modeling, in Current Communications in Molecular Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

5.5. GENERATION OF ANTIBODIES TO SERRATE PROTEINS AND DERIVATIVES THEREOF

According to the invention, Serrate protein, its fragments or other derivatives, or analogs thereof, may be used as an immunogen to generate antibodies which recognize such an immunogen. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. In a specific embodiment, antibodies to human Serrate are produced. In another embodiment, antibodies to the extracellular domain of Serrate are produced. In another embodiment, antibodies to the intracellular domain of Serrate are produced.

Various procedures known in the art may be used for the production of polyclonal antibodies to a Serrate protein or derivative or analog. In a particular embodiment, rabbit polyclonal antibodies to an epitope of the Serrate protein encoded by a sequence depicted in Figure 4, or a subsequence thereof, can be obtained. For the production of antibody, various host animals can be immunized by injection with the native Serrate protein, or a synthetic version, or derivative (e.g., fragment) thereof, including but not limited to rabbits, mice, rats, etc. Various adjuvants may be used to increase the immunological response, depending on the host species, and including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum.

For preparation of monoclonal antibodies directed toward a Serrate protein sequence or analog thereof, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilizing recent technology (PCT/US90/02545). According to the invention, human antibodies may be used and can be obtained by using human hybridomas (Cote et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030) or by transforming human B cells with EBV virus in vitro (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, pp. 77-96). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6851-6855; Neuberger et al., 1984, Nature 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule specific for Serrate together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention.

According to the invention, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce Serrate-specific single chain antibodies. An

additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse et al., 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for Serrate proteins, derivatives, or analogs.

Antibody fragments which contain the idiotype of the molecule can be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g. ELISA (enzyme-linked immunosorbent assay). For example, to select antibodies which recognize a specific domain of a Serrate protein, one may assay generated hybridomas for a product which binds to a Serrate fragment containing such domain. For selection of an antibody specific to human Serrate, one can select on the basis of positive binding to human Serrate and a lack of binding to Drosophila Serrate.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the protein sequences of the invention (e.g., see Section 5.7, infra), e.g., for imaging these proteins, measuring levels thereof in appropriate physiological samples, etc.

5.6. SERRATE DERIVATIVES AND ANALOGS

The invention further relates to derivatives (including but not limited to fragments) and analogs of Serrate proteins.

The production and use of derivatives and analogs related to Serrate are within the scope of the present invention. In a specific embodiment, the derivative or analog is functionally active, i.e., capable of exhibiting one or more functional activities associated with a full-length, wild-type Serrate protein. As one example, such derivatives or analogs which have the desired immunogenicity or antigenicity can be used, for example, in immunoassays, for immunization, for inhibition of Serrate activity, etc. Such molecules which retain, or alternatively inhibit, a desired Serrate property, e.g., binding to Notch or other toporythmic proteins, binding to a cell-surface receptor, can be used as inducers, or inhibitors, respectively, of such property and its physiological correlates. A specific embodiment relates to a Serrate fragment that can be bound by an anti-Serrate antibody but cannot bind to a Notch protein or other toporythmic protein. Derivatives or analogs of Serrate can be tested for the desired activity by procedures known in the art, including but not limited to the assays described in Section 5.7.

In particular, Serrate derivatives can be made by altering Serrate sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as a Serrate gene may be used in the practice of the present invention. These include but are not limited to nucleotide sequences comprising all or portions of

Serrate genes which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change. Likewise, the Serrate derivatives of the invention include, but are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of a Serrate protein including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Derivatives or analogs of Serrate include but are not limited to those peptides which are substantially homologous to Serrate or fragments thereof, or whose encoding nucleic acid is capable of hybridizing to a Serrate nucleic acid sequence.

The Serrate derivatives and analogs of the invention can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level.

For example, the cloned Serrate gene sequence can be modified by any of numerous strategies known in the art (Maniatis, T., 1990, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated in vitro. In the production of the gene encoding a derivative or analog of Serrate, care should be taken to ensure that the modified gene remains within the same translational reading frame as Serrate, uninterrupted by translational stop signals, in the gene region where the desired Serrate activity is encoded.

Additionally, the Serrate-encoding nucleic acid sequence can be mutated in vitro or in vivo, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further in vitro modification. Any technique for mutagenesis known in the art can be used, including but not limited to, in vitro site-directed mutagenesis (Hutchinson, C., et al., 1978, J. Biol. Chem 253:6551), use of TAB® linkers (Pharmacia), etc.

Manipulations of the Serrate sequence may also be made at the protein level. Included within the scope of the invention are Serrate protein fragments or other derivatives or analogs which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody

molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

In addition, analogs and derivatives of Serrate can be chemically synthesized. For example, a peptide corresponding to a portion of a Serrate protein which comprises the desired domain (see Section 5.6.1), or which mediates the desired aggregation activity in vitro, or binding to a receptor, can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the Serrate sequence. Non-classical amino acids include but are not limited to the D-isomers of the common amino acids, α-amino isobutyric acid, 4-aminobutyric acid, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β-alanine, designer amino acids such as β-methyl amino acids, Cα-methyl amino acids, and Nα-methyl amino acids.

In a specific embodiment, the Serrate derivative is a chimeric, or fusion, protein comprising a Serrate protein or fragment thereof fused to a non-Serrate amino acid sequence. In one embodiment, such a chimeric protein is produced by recombinant expression of a nucleic acid encoding the protein (comprising a Serrate-coding sequence joined in-frame to a non-Serrate coding sequence). Such a chimeric product can be made by ligating the

appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the proper coding frame, and expressing the chimeric product by methods commonly known in the art. Alternatively, such a chimeric product may be made by protein synthetic techniques, e.g., by use of a peptide synthesizer. In a specific embodiment, a chimeric nucleic acid encoding a mature Serrate protein with a heterologous signal sequence is expressed such that the chimeric protein is expressed and processed by the cell to the mature Serrate protein. As another example, and not by way of limitation, a recombinant molecule can be constructed according to the invention, comprising coding portions of both Serrate and another toporythmic gene, e.g., Delta. The encoded protein of such a recombinant molecule could exhibit properties associated with both Serrate and Delta and portray a novel profile of biological activities, including agonists as well as antagonists. The primary sequence of Serrate and Delta may also be used to predict tertiary structure of the molecules using computer simulation (Hopp and Woods, 1981, Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828); Serrate/Delta chimeric recombinant genes could be designed in light of correlations between tertiary structure and biological function. Likewise, chimeric genes comprising portions of Serrate fused to any heterologous protein-encoding sequences may be constructed. A specific embodiment relates to a chimeric protein comprising a fragment of Serrate of at least six amino acids. A particular example of the construction and expression of a Notch-Serrate chimera is presented in Section 8 hereof. A particular example of another Serrate fusion protein is presented in Section 7 hereof.

In another specific embodiment, the Serrate derivative is a fragment of Serrate comprising a region of homology with another toporythmic protein. As used herein, a region of a first protein shall be considered "homologous" to a second protein when the amino acid sequence of the region is at least 30% identical or at least 75% either identical or involving conservative changes, when compared to any sequence in the second protein of an equal number of amino acids as the number contained in the region. For example, such a Serrate fragment can comprise one or more regions homologous to Delta, including but not limited to Serrate amino acids 63-73, 124-134, 149-158, 195-206, 214-219, 250-259, or 79-282 (or 79-246, excluding the partial EGF-like repeat) (see Figs. 4, 8), or portions of Serrate of other species most homologous to the foregoing sequences.

Other specific embodiments of derivatives and analogs are described in the subsections below and examples sections infra.

5.6.1. DERIVATIVES OF SERRATE CONTAINING ONE OR MORE DOMAINS OF THE PROTEIN

In a specific embodiment, the invention relates to Serrate derivatives and analogs, in particular Serrate fragments and derivatives of such fragments, that comprise one or more domains of the Serrate protein, including but not limited to the extracellular domain, transmembrane domain, intracellular domain, membrane-associated region, and one or more of the EGF-like repeats (ELR) of the Serrate protein. In particular examples relating to the Drosophila Serrate protein (see example 6), such domains are identified as follows, with reference to Figure 4: extracellular domain, amino acids numbers (AA) 81-541; transmembrane domain, AA 1221-1245;

intracellular domain, AA 1246-1404; membrane-associated region, AA 542-564; ELR (see underscored sequences in Fig. 4).

In a specific embodiment, relating to a Serrate protein of a species other than D. melanogaster, the fragments comprising specific portions of Serrate are those comprising portions in the respective Serrate protein most homologous to specific fragments of the Drosophila Serrate protein. Alternatively, a fragment comprising a domain of a Serrate homolog can be identified by protein analysis methods as described in Section 5.3.2 or 6.

Serrate derivatives which are Serrate fragments and chimeric/fusion proteins are described by way of example in Sections 7 and 8 infra.

5.6.2. DERIVATIVES OF SERRATE THAT MEDIATE BINDING TO TOPORYTHMIC PROTEIN DOMAINS

The invention also provides for Serrate fragments, and analogs or derivatives of such fragments, which mediate binding to toporythmic proteins (and thus are termed herein "adhesive"), and nucleic acid sequences encoding the foregoing.

In a specific embodiment, the adhesive fragment of Serrate is that comprising the portion of Serrate most homologous to about amino acid numbers 85-283 or 79-282 of the Drosophila Serrate sequence (see Figure 4).

The ability to bind to a toporythmic protein (preferably Notch) can be demonstrated by in vitro aggregation assays with cells expressing such a toporythmic protein as well as cells expressing Serrate or a Serrate derivative (See Section 5.7). That is, the ability of a Serrate fragment to bind to a Notch protein can be demonstrated by detecting the ability of the Serrate fragment, when expressed on the

surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell.

The nucleic acid sequences encoding toporythmic proteins or adhesive domains thereof, for use in such assays, can be isolated from human, porcine, bovine, feline, avian, equine, canine, or insect, as well as primate sources and any other species in which homologs of known toporythmic genes can be identified.

5.7. ASSAYS OF SERRATE PROTEINS, DERIVATIVES AND ANALOGS

The functional activity of Serrate proteins, derivatives and analogs can be assayed by various methods.

For example, in one embodiment, where one is assaying for the ability to bind or compete with wild-type Serrate for binding to anti-Serrate antibody, various immunoassays known in the art can be used, including but not limited to competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, in situ immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assyas, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is

labelled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

In another embodiment, where one is assaying for the ability to mediate binding to a toporythmic protein, e.g., Notch, one can carry out an in vitro aggregation assay such as described infra in Section 8.2.1 (see also Fehon et al., 1990, Cell 61:523-534; Rebay et al., 1991, Cell 67:687-699).

In another embodiment, where a receptor for Serrate is identified, receptor binding can be assayed, e.g., by means well-known in the art. In another embodiment, physiological correlates of Serrate binding to cells expressing a Serrate receptor (signal transduction) can be assayed.

In another embodiment, in insect or other model systems, genetic studies can be done to study the phenotypic effect of a Serrate mutant that is a derivative or analog of wild-type Serrate (see Section 6, infra).

Other methods will be known to the skilled artisan and are within the scope of the invention.

6. THE GENE SERRATE ENCODES A PUTATIVE EGF-LIKE TRANSMEMBRANE PROTEIN ESSENTIAL FOR PROPER ECTODERMAL DEVELOPMENT IN DROSOPHILA MELANOGASTER

As described in the example herein (see Fleming et al., 1990, Genes Dev. 4:2188-2201), mutations in the third chromosome gene Serrate are shown to display genetic interactions with specific alleles of the neurogenic locus Notch, which encodes a transmembrane protein with epidermal growth factor homology. The locus Serrate displays a striking phenotypic interaction with a specific Notch allele known to affect postembryonic development. We present

the molecular cloning of Serrate and show that it encodes two coordinately-expressed transcripts from a genomic interval greater than 30 kilobases in length. The deduced protein product of 1404 amino acids contains a single transmembrane domain and 14 epidermal growth factor-like repeats. Whole-mount in situ hybridization analysis revealed complex temporal and spatial patterns of RNA expression consistent with the epidermal and neuronal defects observed in mutant embryos.

We demonstrate that the Serrate locus encodes an essential function, the loss of which results in embryonic lethality brought about by the disruption of both neuronal and epidermal tissues. Serrate is likely to represent an element in a network of interacting molecules operating at the cell surface during the differentiation of certain tissues.

6.1. RESULTS

6.1.1. THE SERRATE AND NOTCH GENES INTERACT PHENOTYPICALLY

In the course of genetic crosses designed to detect interactions between the Notch locus and other genes in Drosophila, a dramatic phenotypic interaction was observed between the Notch allele notchoid (nd) and the third chromosome mutation Serrate (designated Ser^D herein). The recessive nd mutation, which is associated with an amino acid substitution in the intracellular portion of the Notch protein (Xu et al., 1990, Genes Dev. 4:464-475), causes wing notches in the adult (see Fig. 1b; compare to wildtype, Fig. 1a). The Ser^D mutation is dominant and in heterozygous condition produces an adult wing blade very similar to that of nd animals (compare Figures 1b and 1c). The phenotypic interaction seen in nd/Y; Ser^D/+ males is

characterized by loss of anterior and posterior wing margins, as well as loss of distal wing blade tissue. Concomitant with this loss, thickening of the L3 and L5 wing veins is observed (see Fig. 1d).

Even though both the Ser^{D} and nd mutations affect wing blade development, the interaction appears to be synergistic because a novel phenotype is seen, that is, rather than just additive effects. To explore this question of synergy further, we constructed flies carrying genetic duplications of Notch^{+}. Animals carrying an extra copy of Notch^{+} normally exhibit a Confluens phenotype characterized by wing vein thickening. Surprisingly, animals bearing Ser^{D} and an extra copy of Notch^{+} have essentially wild-type wings (Fig. 1e), that is, both the Ser^{D} wing nicking and the Confluens phenotypes are suppressed in this combination. This interaction was noted using both Dp(1;2)51b (a large genetic duplication of 3C1-2; 3D6 including N^{+}) and CosP479BE [(N^{+})(86E5-6), a cosmid construct containing only the N^{+} gene (Ramos et al., 1989, Genetics 123:337:348)].

Because the Ser^{D} mutation is neomorphic, the interactions observed between Ser^{D} and Notch mutations might not be representative of interactions normally occurring between these gene products. We therefore examined the phenotypes of nd males heterozygous for Df(3R)Ser^{+82f24} (nd/Y; Df(3R)Ser^{+82f24}/+). These animals exhibit a significantly increased mutant wing phenotype as compared to nd alone (not shown). Thus, it appears that Notch and Ser^{D} mutually influence each other's phenotypic expression.

6.1.2. GENETIC CHARACTERIZATION OF SERRATE

Previous genetic characterizations have demonstrated that the Ser^{D} mutation maps to the 97F

region of the polytene chromosomes and is neomorphic, producing the dominant wing nicking phenotype shown in Figure 1d (Belt, 1971, Drosophila Inf. Serv. 46:116; P. Lewis, Yale University; unpubl.). The neomorphic nature is demonstrated genetically via the insensitivity of the Ser^D phenotype to the number of wildtype (Ser^+) copies present, that is $Ser^D/+/+$ displays a phenotype similar to $Ser^D/+$ and to Ser^D/Deficiency (P. Lewis, pers. comm.). Flies with only one copy of wild-type Ser^+ (i.e., individuals heterozygous for a wild-type allele over deficiencies) are phenotypically wild-type, demonstrating that reduction of gene product (i.e., haploinsufficiency) is not causing the dominant phenotype. Finally, when the Ser^D mutation is homozygous, viable adults are produced that display a more severe wing phenotype than heterozygous $Ser^D/+$ animals (Figure 1f). Thus, the expression of the Ser^D wing phenotype appears to be directly related to the expression of a mutant or novel gene product rather than to Ser^+ gene dosage.

In an effort to obtain amorphic alleles of Serrate, we used X-ray mutagenesis to produce phenotypic revertants of the dominant mutation (see Section 6.3 for details). All five of the revertants of the Ser^D mutation are lethal when homozygous and, consistent with the deficiency phenotypes, are phenotypically wild-type when heterozygous with a wild-type chromosome. Complementation tests revealed that the $Ser^{revertants}$ are allelic. Moreover, transheterozygotes of nd with two different $Ser^{revertant}$ alleles (nd/Y; $Ser^{rev2-3}/+$ and nd/Y; $Ser^{rev2-11}/+$) exhibit an enhanced mutant wing phenotype as compared to nd mutants, in agreement with the Ser^D-Notch interactions noted previously.

These complementation tests were extended to include another dominant mutation, Beaded of Goldschmidt (Bd^G), which also maps to the 97F region. Heterozygous adults bearing the Bd^G mutation display a wing nicking phenotype that is more severe than that observed in Ser^D heterozygotes (data not shown). Moreover, the Bd^G mutation, unlike Ser^D, is homozygous lethal. Finally, three alleles ($Bd^{43.5}$, $Bd^{862.5}$, and pll^{11}) of a lethal complementation group isolated in K. Anderson's laboratory were shown to be allelic to Bd^G (P. Hecht, unpubl.; a complete listing of the alleles used and their descriptions is provided in Section 6.3). Although transheterozygotes of Ser^D and Bd^G are viable, it is interesting to note that $Df(3R)Ser^{+82f24}$ and most of the $Ser^{revertants}$ fail to complement the Bd^G mutation for viability. The exception is the Ser^{rev2-3} allele, which although homozygous lethal, complements Bd^G. Despite the exceptional Ser^{rev2-3} allele, these results suggest that the Serrate and Beaded mutations are alleles of the same gene (see also below). Consistent with this idea is the fact that $Ser^{revertant}$ and Bd alleles have similar phenotypes (see also below).

6.1.3. PHENOTYPIC CHARACTERIZATION OF SERRATE LETHAL ALLELES

All of the $Ser^{revertants}$, as well as the Bd alleles we tested (Bd^G, $Bd^{43.5}$, and $Bd^{862.5}$), and the $Df(3R)Ser^{+82f24}$ exhibit embryonic lethality. Cuticle preparations of unhatched embryos from heterozygous parents revealed a continuous and complex range of phenotypes: The progeny of a single allele from a single brood included individuals that were nearly wild-type in appearance as well as those lacking the majority of differentiated cuticle. Another commonly

observed defect was the failure of germ band retraction. Although all of the alleles displayed the full range of mutant phenotypes (including Df(3R)Ser^{+82f24}), the proportions of weak or strong phenotypes observed for each individual allele differed. Moreover, three alleles, Ser^{rev2-3}, Ser$^{rev2-11}$ and Ser^{rev5-5}, in heterozygous combination with Df(3R)Ser^{+82f24}, revealed the same range of mutant phenotypes as when homozygous.

Embryos exhibiting weak mutant phenotypes often appear to have completely differentiated cuticular structures yet fail to retract the germ band. These embryos appear "J"-shaped within the egg, with the Filzkörper residing at ~50% egg length on the dorsal surface of the embryo (Fig. 2b). Other weakly affected individuals undergo normal germ band retraction, have faintly pigmented denticle bands and dorsal hairs, and may have "holes" in the cuticle along the length of the embryo. More severe phenotypes are exemplified by embryos with retracted germ bands that lack the cuticle of the entire cephalic regions and sometimes the first and second thoracic segments (Fig. 2c). Other embryos fail to retract the germ band and lack head and thoracic structures, may exhibit twisted germ bands, and/or frequently lack large patches of dorsal or ventral cuticle. Finally, in very severely affected mutant embryos, only a small cuticular patch remains (Fig. 2d). Unlike embryos from a neurogenic mutant, the cuticle that remains in Ser^{D} mutant embryos can be of dorsal or ventral origin.

We examined the nervous system of mutant embryos using anti-horseradish peroxidase (anti-HRP) antibody (Jan and Jan, 1982, Proc. Natl. Acad. Sci. USA 79:2700-2704). Consistent with the cuticular

abnormalities, a corresponding range of defects was observed for homozygous individuals of a given Ser^{revertant} allele. Many individuals exhibited missing commissures or breaks in the longitudinal connectives that run between segmental ganglia (Fig. 2f). In other embryos this disruption was more pronounced such that there appeared to be only one longitudinal connective running most of the length of the embryo; this condition often correlated with improper germ band retraction or twisted germ bands (Fig. 2g). Finally, in the most severely affected individuals, only small clumps of anti-HRP staining "neural" material were present throughout the embryo (Fig. 2h). In no case did we see hypertrophy of the central nervous system.

6.1.4. MOLECULAR CHARACTERIZATION OF SERRATE DNA

In an effort to elucidate the molecular nature of the Serrate gene product, DNA from the 97F region was cloned and characterized. A Drosophila genomic clone, previously isolated on the basis of cross hybridization to the EGF-like domain of the Notch gene (Rothberg et al., 1988, Cell 55:1047-1059), was used as an entry point to initiate a chromosomal walk. From this initial clone, eight recombinant phage spanning ~85 kb of genomic DNA were isolated (see Fig. 3). A BamHI site adjacent to the region of EGF homology was arbitrarily chosen as coordinate position zero.

Genomic Southern blots containing mutant and wild-type DNAs were probed with DNA from the individual phage isolates to detect and localize rearrangement breakpoints that might be associated with the various Serrate alleles. Within the first phage isolate, ø10.2, restriction fragment

polymorphisms were detected on the original Ser^D chromosome. The polymorphism detected with each of three restriction enzymes (EcoR., BamHI, and HindIII) was consistent with an insertion of ~5.5 kb of DNA between map coordinates 0 and -3 (Fig. 3). Subsequent Southern analysis using DNA cloned from Ser^D revealed a repeated DNA sequence, suggesting the presence of a mobile insertional element associated with the mutation. In addition to the insertion, the HindIII site at coordinate -2 has been eliminated in the Ser^D chromosome. Because the parental chromosome from which the Ser^D mutation arose is unavailable, we cannot be certain that the noted polymorphisms are causal to the Ser^D phenotype.

Of the five $Ser^{revertant}$ alleles, three (Ser^{rev2-3}, Ser^{rev5-5}, and $Ser^{rev 6-1}$) appeared cytologically normal and did not exhibit DNA polymorphisms detectable by our Southern analyses. The remaining two revertants, $Ser^{rev2-11}$ and $Ser^{rev 3}$, had polymorphic DNA restriction fragments within the cloned region. $Ser^{rev 2-11}$ is an inversion of polytene bands 97F to 98C. The 97F breakpoint was localized between coordinates +1.5 to +4, within the region of strongest detectable EGF homology (Fig. 3). $Ser^{rev 3}$ is a reciprocal translocation of chromosomes 3R and 2R, with the 97F breakpoint localized between coordinates +15 and +17 (Fig. 3). In situ hybridization of the cloned wild-type genomic DNAs to polytene chromosomes of $Ser^{rev 3}$ and $Ser^{rev 2-11}$ confirmed that the observed DNA polymorphisms represent the 97F breakpoints of these chromosomal rearrangements.

As noted earlier, $Ser^{revertant}$ alleles fail to complement Bd^G, suggesting that the Serrate and Bd mutations are alleles of the same gene. As with the Ser^D mutation, the parental chromosome for the Bd^G

mutation was not available; hence, unambiguous assignment of mutant phenotypes to DNA polymorphisms cannot be made. Cytological observations of the $\underline{Bd}^{G}$ chromosome failed to reveal any visible abnormalities; however, two regions of DNA polymorphism were detectable by Southern analysis. These regions lie between coordinates 0 to +1 and +14 to +17. Investigations of the polymorphism at position 0 to +1 were pursued by cloning the mutant DNA sequences. Preliminary results indicate that the polymorphisms do not result from a small inversion between these two regions but, rather, from a more complex event.

Of the three mutant chromosomes, $\underline{Bd}^{43.5}$, $\underline{Bd}^{862.5}$, and $\underline{pll}^{11}$, only $\underline{pll}^{11}$ was found to have a DNA polymorphism, which was localized between coordinates +17 and +19 (Fig. 3). Genetic and cytological data for the $\underline{pll}^{11}$ mutation suggest the presence of a very small chromosomal aberration within the 97F region (P. Hecht, pers. comm.), and the molecular data are consistent with this observation. Finally, $\underline{T(Y:3)R128}$ is a reciprocal translocation that also breaks within the 97F region (Lindsley et al., 1972, Genetics 71:157-184) and fails to complement $\underline{Bd}^{G}$ (P. Hecht, pers. comm.). The DNA breakpoint for this translocation resides at map coordinates +25 to +28 (Fig. 3). Taken together, these findings strengthen the genetic evidence that Serrate and Bd mutations are alleles of the same gene. In summary, of eleven tested chromosomes containing Serrate or Bd mutation, six were shown to have associated DNA rearrangements within a 30 kb region known to contain EGF homologous sequences.

To examine the structure of the Serrate transcription unit, we probed Northern blots containing 2- to 14-hour embryonic poly(A)$^{+}$ RNA with

the recombinant phages spanning this region (ø10.1, ø1.3 and ø15K; Fig. 3). This analysis revealed the presence of two transcripts of ~5.5 kb and 5.6 kb. We isolated two overlapping cDNA clones, denoted C1 and C3, from an early pupal library (see Section 6.3). Sequence analysis of these cDNAs revealed a perfect overlap of 109 bp for a combined length of 5.6 kb, which is in excellent agreement with the larger of the two transcripts as determined by Northern analysis. Genomic probes unique to the 5' end of C3 only detected the larger 5.6 kb transcript. Thus, the size difference between the 5.5 and 5.6 kb transcripts may represent an alteration in the potential protein coding capacity or an alteration of 5' untranslated sequence. The composite 5.6 kb cDNA confirms that the Serrate transcription unit spans ~30 kb of genomic DNA, encompasses the EGF homologous region, and is interrupted by at least five of the six DNA rearrangements that affect Serrate function (Fig. 3). From Southern analysis, at least two introns are apparent; additional introns are likely but not detectable at this level of resolution.

6.1.5. SERRATE ENCODES A PUTATIVE TRANSMEMBRANE PROTEIN WITH 14 EGF-LIKE REPEATS

The complete nucleotide sequence compiled from the cDNAs C1 and C3 is 5561 bp (see Fig. 4) and agrees with the transcript sizes determined by Northern analysis. Within this sequence there is a single large open reading frame (ORF) of 4329 bp. There are two possible initiator AUG codons at positions 433 and 442. Of these, the second AUG is within a sequence context that agrees with the Drosophila consensus sequence determined for translation initiation [CAAAAUG; (Cavener, 1987, Nucl. Acids Res. 15:1353-1361)]. Predicted codon usage within this ORF is

highly consistent with established Drosophila melanogaster codon preferences (Beachy et al., 1985, Nature 313:545-550). Assuming that translation starts at the second AUG, the Serrate mRNA contains an untranslated leader sequence of at least 441 base pairs, encodes an expected protein product of 1404 amino acids, and terminates with 908 bp of untranslated 3' sequence (Fig. 5a). However, if translation begins at the first AUG, the protein product is 1443 amino acids.

Hydropathy plots revealed three major hydrophobic regions (Fig. 5b; see also Section 6.3). The first, beginning at amino acid 51, is likely to represent a signal peptide sequence; a potential signal cleavage site occurs at amino acid 80. A second hydrophobic domain runs from amino acid 540 to 560. This region does not have a requisite transmembrane structure and is more likely to be a membrane-associated domain. The third hydrophobic domain (amino acids 1220 to 1245) is bounded by hydrophilic residues and is therefore likely to represent a true transmembrane domain.

The most striking structural feature of the predicted protein is the series of EGF-like repeats (see Fig. 5c). There are 14 copies of this motif with an additional partial or degenerate repeat occurring toward the amino terminus (see below). In addition, at least three of these repeats are interrupted by stretches of amino acids. The first interruption (labelled A in Fig. 5c), which occurs in the fourth complete EGF-like repeat (repeats are numbered beginning from the amino terminus), is ~64 amino acids in length and is enriched for serine residues. The second interruption (labelled B in Fig. 5c), occurring in the sixth repeat, is ~44 amino acids long

and has numerous hydrophobic residues. This region represents the putative membrane-associated domain noted earlier. The final interruption (labelled C in Figure 5c), which occurs in the tenth repeat and is 29 amino acids in length, has an unusual run of threonines [$Thr_{(9)}$ Ala $Thr_{(3)}$].

Within the amino-terminal region of the Serrate protein, considerable structural homology (darkly-shaded region in Fig. 5c) is observed with the main protein product of the Delta locus (Vässin et al., 1987, EMBO J. 6:3431-3440; Kopczynski et al., 1988, Genes Dev. 2:1723-1735). Near the signal peptides for both of these molecules there lies a stretch of ~210 conserved amino acids. Within the first 165 amino acids, there is ~32% identity, which increases to greater than 50% for the remaining 45 amino acids. The latter region corresponds to the partial EGF-like repeat (designated PR in Fig. 5c), which lacks a cysteine residue but retains the other characteristic cysteines and conserved amino acids typically found in the remaining EGF-like repeats. The homology between Serrate and Delta extends beyond these amino-terminal regions, since both of these proteins contain EGF-like repeats.

In addition to the extracellular EGF-like sequences, the predicted Serrate protein contains a small intracellular domain of ~160 amino acids. The internal domain does not contain any significant known structural homologies, although there are numerous potential sites for phosphorylation (Those identified in the putative intracellular region by the SITES program were at amino acid positions 1283, 1292, 1297, 1349, 1365, 1371, 1389, and 1390).

6.1.6. EXPRESSION OF SERRATE RNA

Northern analysis of developmentally staged RNAs revealed that the majority of Serrate expression is represented by two coordinately regulated transcripts of 5.5 kb and 5.6 kb, which first appear 4 to 8 hours into embryogenesis (Fig. 6). These transcripts show peak expression between 8 and 12 hours of embryogenesis and diminish thereafter; however, they continue to be readily detectable throughout development except for the adult stages (Fig. 6). In addition to these major transcripts, a smaller (3.4 kb) transcript is expressed transiently between 2 and 4 hours of embryogenesis (Fig. 6).

We undertook an analysis of the spatial distribution of RNA transcripts from the Serrate locus in order to identify regions of the embryo that may require Serrate function. Using the whole mount in situ method (Tautz and Pfeifle, 1989, Chromosoma 98:81-85) and employing nonradioactive probes that hybridize to both the 5.5 kb and 5.6 kb transcripts, we found that Serrate mRNA accumulates in a dynamic pattern beginning from mid-embryogenesis (late stage 10) and persisting until the latest stages examined (stage 16); (embryonic stages are those of Campos-Ortega and Hartenstein, 1985, The Embryonic Development of Drosophila Melanogaster, Springer-Verlag, Berlin). Because the tissue distribution of the two transcripts may be independently regulated, we note that the observed RNA localizations may represent a composite for both transcripts. We also note the possibility of a low level of Serrate RNA in the yolk of pre-gastrulation embryos because faint staining of the yolk was observed consistently. Although this staining was never observed with control probes (see Section 6.3), the presence of yolk staining is known to be a common artifact of the whole-mount in situ

technique (Ashburner, 1989, Drosophila - A laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). However, if this observation is not artifactual, the observed staining may correspond to the expression of the transient 3.4 kb RNA species observed by the Northern analysis of this same developmental stage.

Initial cellular localization was seen in late-stage 10 embryos and consisted of a ring of cells in the foregut. The foregut is formed by the invagination of the stomodeum (the initial event of stage 10); thus, the foregut is actually derived from ectodermal tissue. Shortly thereafter, a bilateral patch of expressing cells appeared in the anterior-most portion of the head, the presumptive clypeolabrum (Fig. 7a). Additional areas of expression appeared abruptly at the end of stage 10 in a group of cells on the lateral edge of abdominal segment 8, followed by cells near the proctodeum and lateral epidermis of abdominal segment 9 (Fig. 7b). Later, during stage 11, expression was detected within cells located at the junction between the labial and maxillary lobes and within cells located near the tracheal pit of the first thoracic segment. The expression pattern progressed to include a group of lateral epidermal cells located between the tracheal pits in each of the thoracic and abdominal segments (Fig. 7c). In addition, each abdominal segment displayed a cluster of cells on either side of the ventral midline.

During germ band retraction (stage 12), the lateral epidermal cell patches broadened to form stripes that lie in the middle of each segment. A portion of these cells appeared to coalesce into an internal longitudinal stripe that was coincident with the developing tracheae (see Fig. 7 e, f, g, and h).

The cells that remained on the surface extended dorsally and ventrally forming a zig-zag shaped pattern (Fig. 7g, arrows). This surface expression in the thoracic segments was wider, more intense, and extended further dorsally and ventrally than in the abdominal segments (Fig. 7g). Later in embryogenesis (stages 14 and 15) the surface epidermal expression, with the exception of the first thoracic segment, diminished relative to the tracheal expression. Later, intense expression was observed in what appeared to be ectodermal invaginations located dorsolaterally on the thoracic segments (Fig. 7j). These pockets of cells may correspond to primordia of imaginal discs; in the first thoracic segment they appeared to be closely associated with opening of the anterior spiracle.

Coincident with the lateral expression, another segmentally reiterated pattern evolved in the ventral epidermis of the trunk. In the extended germ band embryo, this pattern, which consisted of stripes of expressing cells near the anterior border of the abdominal segments, lay out of register with the corresponding lateral expression (Fig. 7c). The pattern in the thorax contrasted with that in the abdomen and consisted of only small clusters of expressing cells in the latero-ventral region (see Fig. 7f and h). The ventral expression was quite intense through stage 13 and dissipated thereafter (Fig. 7l).

Serrate expression was also observed in the ectodermally-derived portions of the gut. The earliest expression was evident in the foregut and persisted throughout embryonic development (Fig. 7a). During germband retraction, a tightly defined, intensely expressing ring of cells lay at the junction

with the anterior midgut. The proventriculus develops from this area; however, expression was limited to the ectodermally-derived portion of this composite structure (King, 1988, J. Morph. 196:253-282). Hindgut expression, though appearing later than foregut expression, occurred at an analogous position, that is, where ectoderm meets endoderm. The initial expression in the hindgut was seen at the time of germ band retraction (stage 12) as a wide band of cells where the Malphigian tubules were forming, but never included the tubules themselves. Later still (stage 14), an additional ring of expression appeared in the hindgut approximately mid-way between the insertion point of the Malphigian tubules and the proctodeum (Fig. 7h). Expression at the posterior-most end of the embryo, near the proctodeal opening, initiated early (stage 11) (Fig. 7b). This expression within the telson remained at high levels throughout embryonic development, eventually forming a ring of cells around the presumptive anal pads (Fig. 7l).

Within the head region, Serrate expression was temporally and spatially dynamic. The earliest expression occurred in the presumptive clypeolabrum (stage 10; Fig. 7a) and became broader and more intense as development proceeded. Early expression between the labial and maxillary lobes increased along their borders, and expression was also seen in the anterior of the mandibular lobe during stage 12 (Fig. 7d and e). In addition, expression was now observed in the hypopharyngeal region, just posterior to the stomodeum, and at the base of the labial lobes in an area encompassing the salivary gland duct opening (Fig. 7d). There was also low level expression in the dorsal procephalic epidermal region (not shown). By the end of germ band retraction (stage 13), expression

encompassed the entire mandibular lobe. As a consequence of the cellular movements associated with head involution (stages 14-16), the expressing cells of the clypeolabrum, hypopharynx and labial lobes combined to form the pharynx. Prior expression in the area of the salivary gland placodes was now limited to the ducts of the developing salivary gland (Fig. 7i). The maxillary and mandibular lobes, which have moved to the anterior-most region of the embryo, expressed intensely at this time (Fig. 7j).

Serrate expression in the central nervous system (CNS) was apparent during stage 12 as a segmentally-reiterated array of single cells along the lateral edge of the ventral nerve cord and within the supraesophogeal ganglia (brain hemispheres). By the end of germ band retraction (stage 13), there were now two cells that appeared to express in each hemisegment of the ventral nerve cord (not shown). However, by stage 15, ventral nerve cord expression was again limited to a single cell per hemisegment (Fig. 7i) while expression in the brain hemispheres remained unchanged (Fig. 7k).

In summary, there are a wide array of tissues that express Serrate mRNA, and the expression pattern is tightly regulated both temporally and spatially. In addition, it should be stressed that at the present level of resolution, Serrate expression appears to be restricted exclusively to cells of ectodermal origin.

6.2. DISCUSSION

Unlike Notch and Delta, the fourteen EGF repeats of Serrate are not completely contiguous. At least three of these repeats contain sizeable interruptions consisting of insertions of long stretches of amino acids. Similarly, interruptions

have been noted in two of the thirty EGF-like repeats of the Drosophila gene crumbs (Tepass et al., 1990, Cell 61:787-799). In Serrate, the interruption that occurs in the sixth repeat is particularly intriguing because it consists largely of hydrophobic amino acids. Although hydropathy plots indicate that this region does not conform to known transmembrane regions, it could represent a membrane-associated domain that serves to "tie" the protein back to the membrane. The interruption in the tenth repeat is also unusual in that it bears a stretch of threonines [$Thr_{(9)}$Ala $Thr_{(3)}$]. A similar motif of thirteen contiguous threonine residues is found in the glycoprotein glutactin, a basement membrane protein of Drosophila (Olson et al., 1990, EMBO J. 9:1219-1227).

If the observed genetic interactions between Notch and Serrate had been only with the original Ser^{D} allele, it could have been argued that this neomorphic mutation is allowing two functionally disparate but structurally similar molecules to interact out of their normal contexts. But because we observe genetic interactions with other Serrate alleles, it is likely that we are observing a manifestation of normal Serrate-Notch interactions.

We have shown that phenotypic revertants of Ser^{D} behave genetically in a similar fashion to known deficiencies for the locus; that is, they are homozygous lethal during embryogenesis and completely recessive as heterozygotes. We also gathered evidence indicating that the mutation Bd^{G}, which was thought to belong to a distinct complementation group, may in fact be an allele of Serrate.

The embryonic lethal phenotypes of Ser^{rev2-3}, $Ser^{rev2-11}$, and Ser^{rev5-5}, which are essentially indistinguishable from one another, appear unchanged

when in homozygous or hemizygous condition. This latter result genetically defines these alleles genetically as amorphic. However, since the $\underline{Ser}^{rev2-3}$ allele complements the $\underline{Bd}^{G}$ mutation, the $\underline{Ser}^{rev2-3}$ mutation is probably not a protein null allele.

Consistent with the defects observed in the cuticle and nervous system of $\underline{Ser}^{-}$ embryos, Serrate transcripts are localized in complex patterns within these tissues. The abundant and widespread expression of Serrate transcripts in the segments that make up the embryonic head and thorax correlates well with the lack of embryonic head and thoracic structures commonly seen in $\underline{Ser}^{-}$ embryos. Likewise, the pattern of Serrate expression in the ventral epidermis of the abdominal segments correlates with the frequently absent or improperly formed denticles. Although Serrate is expressed in a small number of cells within the CNS, the gross morphological defects observed in the CNS of $\underline{Ser}^{-}$ embryos may reflect contributions from two components. The first is the loss of Serrate CNS expression itself, and the second may be a consequence of mechanical stresses (e.g., lack of germ band retraction) imposed by an improperly differentiating epidermis.

In the course of examining the embryonic phenotypes associated with Serrate lethal mutations, we noticed their similarity to those produced by several alleles of the gene coding for the Drosophila EGF receptor homolog known as DER, faint little ball or torpedo (Livneh et al., 1985, Cell 40:599-607; Price et al., 1989, Cell 56:1085-1092; Schejter and Shilo, 1989, Cell 56:1093-1104).

6.3. MATERIALS AND METHODS

6.3.1. DROSOPHILA CULTURES AND STRAINS

Cultures were maintained on standard cornmeal/molasses/agar Drosophila medium supplemented with active dry yeast and were raised at 25°C. The red Ser^D, $Df(3R)Ser^{+82f24}$, and Bd^G chromosomes were obtained from Peter Lewis. The red Ser^D chromosome was maintained in homozygous condition. The mutations pll^{11}, $Bd^{862.5}$, and $Bd^{43.5}$ were generously provided by Kathryn Anderson. The Notch duplication CosP479 is an ~40 kb P-element cosmid construct inserted into the third chromosome (Ramos et al., 1989, Genetics 123:337-348). Other mutations and chromosomes have been described previously (Lindsley and Grell, 1968, Genetic variations of Drosophila melanogaster, Carnegie Inst. Wash. Publ. 627).

6.3.2. MUTAGENESIS

Males aged 3-7 days and homozygous for the red Ser^D chromosome were irradiated with approximately 4500 R (150 kV, 5 mA, 9.2 min exposure; Torrex 150 Source, Torr X-Ray Corp.) and mated immediately to $C(1)A;y/y^2Y611$ or $C(1)Dx;yf/y^2Y611$ virgin females. The F_1 males were scored for the absence of the Ser^D wing phenotype and mated to Gl^{p1-3} fz red e/Tm2, red e virgin females to establish balanced Ser^{rev}/Tm2, red e stocks.

Mutations used in this study are shown in Table 1.

TABLE I

Mutation	Origin	Description
Ser^{D}	Spontaneous; information (Lindsley and Grell, 1968, Carnegie Inst. Wash. Publ. 627)	heterozygous dominant wing phenotype, homozygous viable; cytologically normal
Bd^{G}	recovered among heat-treated flies (Gottschewski, 1935, Dros. Inf. Serv. 4:14,16)	heterozygous dominant wing phenotype homozygous lethal, cytologically normal
Ser^{rev3}	X-ray (this study)	homozygous lethal; reciprocal translocation of 3R (97F) to 2R (57)
Ser^{rev2-3}	X-ray (this study)	Homozygous lethal; cytologically normal
$Ser^{rev2-11}$	X-ray (this study)	homozygous lethal; inversion of 97F to 98C
Ser^{rev5-5}	X-ray (this study)	homozygous lethal; cytologically normal
Ser^{rev6-1}	X-ray (this study)	homozygous lethal; cytologically normal
$Bd^{43.5}$	EMS (K. Anderson, unpubl.)	homozygous lethal; cytologically normal
$Bd^{862.5}$	EMS (K. Anderson, unpubl.)	homozygous lethal; cytologically normal
pll^{11}	EMS (K. Anderson, unpubl.)	homozygous lethal; possible small inversion within the 97F interval
T(Y:3)R128	X-ray (Lindsley et al., 1972, Genetics 71:157-184)	homozygous lethal; reciprocal translocation of 3R (97F) to Y short arm
$Df(3R)Ser^{+82f24}$	X-ray (P. Lewis, unpubl.)	deficiency for chromosome bands 97D to 97F-98A1

6.3.3. EMBRYONIC PHENOTYPE ANALYSIS

Cuticle preparations were according to the protocol of Wieschaus and Nüsslein-Volhard (1986, in Drosophila. A Practical Approach, (ed. D.B. Roberts), IRL Press, Oxford, pp. 199-227) on embryos aged for a minimum of 24 hours at 25°C. Anti-horseradish peroxidase antibody staining of the embryonic nervous system (Jan and Jan, 1982, Proc. Natl. Acad. Sci. USA 79:2700-2704) was carried out using fluorescein-conjugated antibody (Cappel) as described in Preiss et al. (1988, EMBO J. 7:3917-3927). CNS preparations of torpedo^{2C82} were used for comparison studies.

6.3.4. ISOLATION OF NUCLEIC ACIDS

Genomic DNA was isolated as described in Pirrotta et al. (1983, EMBO J. 2:927-934). Restriction enzyme cleavage, agarose gel electrophoresis, capillary transfer to nitrocellulose and hybridization conditions were carried out according to standard procedures. DNA probes labeled with ^{32}P were prepared by random oligonucleotide priming, as described in Feinberg and Vogelstein (1983, Anal. Biochem. 132:6-13). Stage-specific total RNAs from a Canton-S strain were extracted in guanidinium thiocyanate essentially as described in Chirgwin et al. (1979, Biochem. 18:5294-5299). Pupal and adult RNAs were generously provided by A. Preiss (Preiss et al., 1988, EMBO J. 7:3917-3927). Poly $(A)^+$ RNA was selected by serial passage over oligo(dT)-cellulose (Stratagene) according to Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York) and stored in ethanol. RNA was size fractionated in formaldehyde agarose gels and blotted onto Nytran membrane (Schleicher & Schuell) by

capillary transfer. RNA was fixed to the membrane via UV crosslinking.

Two Drosophila genomic phage libraries (Preiss et al., 1985, Nature 313:27-32; R. Karess, unpubl.) were screened and recombinant clones were isolated as described in Benton and Davis (1977, Science 196:180-182). cDNAs in λgt10 were isolated from the early pupal library of Poole et al. (1985, Cell 40:37-43). We isolated the C1 cDNA using the genomic EGF-like sequences from coordinates +1.5 to +4 (Fig. 3) as probe. Subsequently, we isolated the C3 cDNA using the 5' 700 bp terminal fragment of the C1 cDNA as probe.

6.3.5. SEQUENCING AND ANALYSIS

The EcoRI cDNA inserts from λgt10 were subcloned directly into Bluescript KS+ and KS- vectors (Stratagene). Single-stranded DNAs were produced according to the manufacturer's instructions. Both strands of the cDNAs were sequenced using the dideoxynucleotide chain-termination procedure (Sanger, et al., 1977, Proc. Natl. Acad. Sci. USA 74:5463-5467) using the Sequenase kit (U. S. Biochemical). Sequence was obtained using the M13 and reverse primers for these vectors. Additional sequence was obtained by generating internal deletions through the use of restriction sites within the Bluescript polylinker and the cDNA inserts. The remaining cDNA sequences that were not accessible by these methods were obtained by using synthetic primers (Research Genetics) complementary to the end of a previously determined sequence.

Sequences were entered by sonic digitizer and overlapping sequence compilation; manipulation, translation, and secondary structure prediction were accomplished by using the Intelligenetics PC-GENE.

Open reading frame prediction and plotting were performed using the University of Wisconsin program CODONPREFERENCE (Gribshov et al., 1984, Nucl. Acids Res. 12:539-549). The SITES program (PCGENE) was used to predict the location of the signal sequence, transmembrane domain, EGF-like repeats, and phosphorylation sites.

6.3.6. WHOLE MOUNT IN SITU PROCEDURE

A modification of the whole-mount in situ procedure of D. Tautz (Procedure 84a in Ashburner, 1989, Drosophila: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York) was used. The differences were as follows: Proteinase K (Boehringer-Mannheim) treatment was 10 to 14 minutes; 100 μl (rather than 10 μl) of boiled probe was used; after washing the embryos with 1:4 hybridization buffer to PBT, they were washed twice in PBT for 20 minutes, and then twice in 1X PBS, 0.1% BSA (globin free, Sigma), 0.2% Triton-X100 for 20 minutes; the antibody treatment was done in the same PBS, BSA, Triton solution at 4°C overnight; the embryos were washed four times in the PBS, BSA, Triton solution at room temperature; after the alkaline phosphatase reaction, embryos were dehydrated twice in 70% and 100% ethanol and then cleared in xylenes; the embryos were mounted in Permount (Sigma). Dissected embryos were rehydrated, dissected in PBT, and mounted in 90% glycerol [10% Tris-HCl at pH 8.0, with 0.5% n-propyl-galate (wt/vol; Sigma)].

The probe was made by runoff of a PCR reaction in 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% (wt/vol) gelatin, 0.2 mM dATP, 0.2 mM dCTP, 0.2 mM dGTP, 0.15 mM dTTP, and 0.07 mM digoxigenin-11-dUTP (Boehringer Mannheim) using 150 ng of custom

synthesized primer and approximately 400 ng of linearized DNA. Probe was synthesized from cDNA coordinates 4826 to 3854; the opposite strand constituted the control probe and was synthesized from coordinates 4458 to 5015 (refer to Fig. 6). The conditions for the PCR thermal cycler were 95°C for 45 seconds, 55°C for 30 seconds, and 72°C for 1 minute, which were run for 30 cycles. The probe was ethanol precipitated twice and resuspended in 300 μl of hybridization solution.

7. EXPRESSION OF A SERRATE FRAGMENT AS A FUSION PROTEIN AND PRODUCTION OF ANTIBODIES THERETO

Mouse anti-Serrate polyclonal antisera were made as follows: A BamHI fragment encoding amino acids 78-425 (Fig. 4) was subcloned into the pGEX-1 expression vector (Smith and Johnson, 1988, Gene 67:31-40). Fusion proteins were purified on glutathione-agarose beads (SIGMA), and injected into mice for antibody production. Mouse antisera were precipitated with 50% $(NH_4)_2SO_4$ and resuspended in PBS (150 mM NaCl, 14 mM Na_2HPO_4, 6 mM NaH_2PO_4) with 0.02% NaN_3.

8. EXPRESSION OF SERRATE AND A FRAGMENT AND A CHIMERIC DERIVATIVE THEREOF; IDENTIFICATION OF A NOTCH-BINDING DOMAIN

We describe herein the recombinant expression of Serrate, of a deletion construct (fragment) thereof, and of a chimeric Notch-Serrate fragment, and show that the full-length Serrate and the chimeric derivative are capable of binding to Notch in vitro.

8.1. EXPRESSION OF SERRATE AND OF DERIVATIVES THEREOF

For the Serrate expression construct, a synthetic primer containing an artificial BamHI site immediately 5' to the initiator AUG at position 442 (all sequence numbers are according to Fleming et al., 1990, Genes & Dev. 4:2188-2201) and homologous through position 464, was used in conjunction with a second primer from position 681-698 to generate a DNA fragment of ~260 base pairs. This fragment was cut with BamHI and KpnI (position 571) and ligated into Bluescript KS+ (Stratagene). This construct, BTSer5'PCR, was checked by sequencing, then cut with KpnI. The Serrate KpnI fragment (571 - 2981) was inserted and the proper orientation selected, to generate BTSer5'PCR-Kpn. The 5' SacII fragment of BTSer5'PCR-Kpn (SacII sites in Bluescript polylinker and in Serrate (1199)) was isolated and used to replace the 5' SacII fragment of cDNA C1 (Fleming et al., 1990, Genes & Dev. 4:2188-2201), thus regenerating the full length Serrate cDNA minus the 5' untranslated regions. This insert was isolated by a SalI and partial BamHI digestion and shuttled into the BamHI and SalI sites of the metallothionein promoter vector pRmHa-3 (Bunch et al., 1988, Nucl. Acids. Res. 16:1043-1061) to generate the final expression construct, Ser-mtn.

A Serrate deletion expression construct was also made, in which nucleotides 672-1293 (encoding amino acids 77-284) (Figs. 4, 8, 9) were deleted. This deletion construct was made as follows: The Ser-mtn construct was digested with EcoRV, which cuts at nucleotide 672, and with SfiI, which cuts at nucleotide 4073. The linearized vector, lacking the EcoRV-SfiI (672-4073) fragment, was isolated. Plasmid

SerFL was then digested with NdeI, which cuts at nucleotide 1289, and treated with mung bean nuclease resulting in the "trimming back" of four bases. The resulting SerFL fragment was then digested with SfiI which cuts at base 4073, and the resulting 1293-4073 fragment was isolated and ligated into the EcoRV-SfiI vector isolated above.

In addition, a Notch-Serrate chimeric construct was made using a clone consisting of Drosophila Notch cDNA with a deletion of all the Notch EGF-like repeats ("ΔEGF") (see copending application Serial No. to be assigned, filed November 14, 1991 by Artavanis-Tsakonas et al.; Rebay et al., 1991, Cell 67:687-699 (Fig. 13, construct no. 25)). An N-terminal region of Serrate with homology to Delta and including the Serrate EGF-like repeats (Serrate nucleotide numbers 676-1287, encoding amino acids 79-282; Figs. 8, 9) was placed into the ΔEGF deletion of Notch.

The above constructs were expressed in Drosophila S2 cells. The S2 cell line (Schneider, 1972, J. Embryol. Exp. Morph. 27, 353-365) was grown in M3 medium (prepared by Hazleton Co.) supplemented with 2.5 mg/ml Bacto-Peptone (Difco), 1 mg/ml TC Yeastolate (Difco), 11% heat-inactivated fetal calf serum (FCS) (Hyclone), and 100 U/ml penicillin-100 μg/ml streptomycin-0.25 μg/ml fungizone (Hazleton). Cells growing in log phase at ~2 x 10^6 cells/ml were transfected with 20 μg of DNA-calcium phosphate coprecipitate in 1 ml per 5 ml of culture as previously described (Wigler et al., 1979, Proc. Natl. Acad. Sci. USA 78, 1373-1376), with the exception that BES buffer (SIGMA) was used in place of HEPES buffer (Chen and Okayama, 1987, Mol. Cell. Biol. 7, 2745-2752). After 16-18 hr, cells were transferred to conical centrifuge tubes, pelleted in a clinical

centrifuge at full speed for 30 seconds, rinsed once with 1/4 volume of fresh complete medium, resuspended in their original volume of complete medium, and returned to the original flask. Transfected cells were then allowed to recover for 24 hr before induction. Expression from the metallothionein constructs was induced by the addition of $CuSO_4$ to 0.7 mM.

8.2. AGGREGATION ASSAYS FOR BINDING TO NOTCH

8.2.1. METHODS

Two types of aggregation assays were used. In the first assay, a total of 3 ml of cells (5-10 x 10^6 cells/ml) was placed in a 25 ml Erlenmeyer flask and rotated at 40-50 rpm on a rotary shaker for 24-48 hr at room temperature. For these experiments, cells were mixed 1-4 hr after induction began and induction was continued throughout the aggregation period. In the second assay, ~0.6 ml of cells were placed in a 0.6 ml Eppendorf tube (leaving a small bubble) after an overnight induction (12-16 hr) at room temperature and rocked gently for 1-2 hr at 4°C. Ca^{2+} dependence experiments were performed using the latter assay. For Ca^{2+} dependence experiments, cells were first collected and rinsed in balanced saline solution (BSS) with 11% FCS (BSS-FCS; FCS was dialyzed against 0.9% NaCl, 5 mM Tris [pH 7.5]) or in Ca^{2+} free BSS-FCS containing 10 mM EGTA (Snow et al., 1989, Cell 59: 313-323) and then resuspended in the same medium at the original volume.

For viewing by immunofluorescence, cells were collected by centrifugation (3000 rpm for 20 seconds in an Eppendorf microcentrifuge) and fixed in 0.6 ml Eppendorf tubes with 0.5 ml of freshly made 2% paraformaldehyde in PBS for 10 min at room

temperature. After fixing, cells were collected by centrifugation, rinsed twice in PBS, and stained for 1 hr in primary antibody in PBS with 0.1% saponin (SIGMA) and 1% normal goat serum (Pocono Rabbit Farm, Canadensis, PA). Sera were appropriately diluted (e.g., 1:1000) for this step. Cells were then rinsed once in PBS and stained for 1 hr in specific secondary antibodies (double-labeling grade goat anti-rabbit and goat anti-mouse), in PBS-saponin-normal goat serum. After this incubation, cells were rinsed twice in PBS and mounted on slides in 90% glycerol, 10% 1 M Tris (pH 8.0), and 0.5% n-propyl gallate. Cells were viewed under epifluorescence on a Leitz Orthoplan 2 microscope.

Confocal micrographs were taken using the Bio-Rad MRC 500 system connected to a Zeiss Axiovert compound microscope. Images were collected using the BHS and GHS filter sets, aligned using the ALIGN program, and merged using MERGE. Fluorescent bleed-through from the green into the red channel was reduced using the BLEED program (all software provided by Bio-Rad). Photographs were obtained directly from the computer monitor using Kodak Ektar 125 film.

Notch-expressing cells for the assays were obtained similarly, using metallothionein promoter-driven plasmid constructions containing D. melanogaster Notch (see copending application Serial No. to be assigned, filed November 14, 1991 by Artavanis-Tsakonas et al.; Fehon et al., 1990, Cell 61:523-534; Rebay et al., 1991, Cell 67:687-699).

8.2.2. RESULTS

We found that Serrate expressing cells adhere to Notch expressing cells in a calcium dependent manner (see also Rebay et al., 1991, Cell 67:687-699).

However, unlike Delta, under the experimental conditions tested, Serrate did not appear to interact homotypically. In addition, we detect no interactions between Serrate and Delta. It is possible that such interactions do occur, but at an affinity such that they are below the level of detection in our assay system.

We have tested a subset of our Notch deletion constructs to map the Serrate-binding domain and have found that Notch EGF-like repeats 11 and 12, in addition to binding to Delta, also mediate interactions with Serrate. In addition, the Serrate-binding function of these repeats also appears to have been conserved in the corresponding two EGF repeats of Xenopus Notch (construct #33ΔC1a+XEGF(10-13); see Rebay et al., supra).

We were also able to define the Serrate region which is essential for the Notch/Serrate aggregation. Deleting nucleotides 672-1293 (i.e. amino acids 77-284) eliminated the ability of the Serrate protein to aggregate with Notch. While both cells expressing Notch and cells expressing the Serrate fragments were detected by immunofluorescence with anti-Notch and anti-Serrate antibodies, respectively, these cells did not co-aggregate.

Aggregation assays with cells expressing Notch and cells expressing the chimeric ΔEGF Notch-Serrate construct showed binding between Notch and the chimeric construct. These experiments thus demonstrated that a fragment of Serrate consisting of amino acids 79-282 (see SEQ ID NO:2) is capable of mediating binding to Notch. Similar experiments with Delta from the laboratory of M. Muskavitch (personal communication) have demonstrated that the homologous region of Delta (without the partial EGF-like repeat)

was sufficient to mediate Notch-Delta binding. Therefore, it is likely that the partial EGF-like repeat of Serrate is not essential for this binding to occur.

Work in our laboratory has shown that Notch and Delta proteins interact directly at the molecular level (Fehon et al., 1990, Cell 61:523-534; copending U.S. patent applications serial no. 07/695,189 filed May 3, 1991 and serial no. to be assigned, filed November 14, 1991, by Artavanis-Tsakonas, et al.; collectively incorporated by reference herein in their entireties), as demonstrated by the specific binding of Notch-expressing cells to Delta-expressing cells in vitro. We have also shown that EGF-like repeats repeats 11 and 12 of Notch are required and sufficient for Notch-Delta-mediated aggregation, and that Delta participates in heterotypic (Delta-Notch) and homotypic (Delta-Delta) interactions mediated by its amino-terminus (id.). Thus, it is conceivable that the Serrate and Delta proteins compete for binding with the Notch protein. Such interplay could underlie the genetic interactions observed between Notch and Serrate.

Notch and Serrate appeared to aggregate less efficiently than Notch and Delta, perhaps because the Notch-Serrate interaction is weaker. For example, when scoring Notch-Delta aggregates, we detect ~40% of all Notch expressing cells in clusters with Delta expressing cells and ~40% of all Delta expressing cells in contact with Notch expressing cells. For Notch-Serrate, we find only ~20% of all Notch expressing cells and ~15% of all Serrate expressing cells in aggregates. For the various Notch deletion constructs tested, we consistently detect a reduction in the amount of aggregation between Notch and Serrate

as compared to the corresponding Notch-Delta levels, with the possible exception of two constructs which exhibit severely reduced levels of aggregation even with Delta. One trivial explanation for this reduced amount of aggregation could be that our Serrate construct simply does not express as much protein at the cell surface as the Delta construct, thereby diminishing the strength of the interaction. Alternatively, the difference in strength of interaction may indicate a fundamental functional difference between Notch-Delta and Notch-Serrate interactions that may be significant in vivo.

9. THE CLONING, SEQUENCING, AND EXPRESSION OF HUMAN SERRATE

Clones for the human Serrate sequence are obtained as described below.

The polymerase chain reaction (PCR) was used to amplify DNA from a 17-18 week human fetal brain cDNA library in the Lambda Zap II vector (Stratagene). Degenerate primers used in this reaction were designed based on amino-terminal regions of homology between Drosophila Serrate and Drosophila Delta (see Fig. 9). Synthetic oligonucleotide primers FLE1 and FLE4R (shown in Fig. 9) were used in PCR to amplify Serrate-homologous fragments in the cDNA library. The PCR reaction products were subjected to agarose gel electrophoresis, resulting in the detection of two amplified DNA fragments from the cDNA library. The fragments are then each cloned into a plasmid for production of quantities thereof, using the TA cloning kit (Invitrogen).

The Serrate-homologous fragments amplified and obtained in this manner are then sequenced at least in part, by use of Sequenase® (U.S. Biochemical Corp.), to confirm the identity of the fragments as Serrate

homologs. Upon such confirmation, the fragments are then used as probes with which to screen the same cDNA library for human Serrate clones. The isolated phage λ clones are converted to plasmids via the manufacturer's instructions, yielding the Serrate-homologous fragments as cDNA inserts in the EcoRI site of pBluescript SK⁻ (Stratagene). The host E. coli strain is XL1-Blue (see Sambrook et al., 1989, Molecular Cloning Press, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, p. A12).

The sequence of human Serrate contained in the cDNA clone(s) is determined (by use of Sequenase®, U.S. Biochemical Corp.).

Expression constructs are made using the isolated clone(s). The clone is excised from its vector as an EcoRI restriction fragment(s) and subcloned into the EcoRI restriction site of each of the three pGEX vectors (Glutathione S-Transferase expression vectors; Smith and Johnson, 1988, Gene 7: 31-40). This allows for the expression of the human Serrate protein product from the subclone in the correct reading frame.

10. DEPOSIT OF MICROORGANISM

Bacteria strain XL1-Blue containing plasmid SerFL, containing an EcoRI fragment encoding a full-length Drosophila Serrate, was deposited on December 11, 1991 with the American Type Culture Collection, 1201 Parklawn Drive, Rockville, Maryland 20852, under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedures, and assigned Accession No. _____.

The present invention is not to be limited in scope by the microorganism deposited or the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Fleming, Robert J.
Artavanis-Tsakonas, Spyridon

(ii) TITLE OF INVENTION: Nucleotide And Protein Sequences Of The Serrate Gene And Methods Based Thereon

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Pennie & Edmonds
(B) STREET: 1155 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10036-2711

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Misrock, S. Leslie
(B) REGISTRATION NUMBER: 18,872
(C) REFERENCE/DOCKET NUMBER: 7326-005

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 212 790-9090
(B) TELEFAX: 212 8698864/9741
(C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5561 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 442..4653

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCGAGTCGAG CGCCGTGCTT CGAGCGGTGA TGAGCCCCTT TTCTGTCAAC GCTAAAGATC 60

TACAAAACAT CAGCGCCTAT CAAGTGGAAG TGTCAAGTGT GAACAAAACA AAAACGAGAG 120

AAGCACATAC TAAGGTCCAT ATAAATAATA AATAATAATT GTGTGTGATA ACAACATTAT 180

CCAAACAAAA CCAAACAAAA CGAAGGCAAA GTGGAGAAAA TGATACAGCA TCCAGAGTAC 240

GGCCGTTATT CAGCTATCCA GAGCAAGTGT AGTGTGGCAA AATAGAAACA AACAAAGGCA 300

```
CCAAAATCTG CATACATGGG CTAATTAAGG CTGCCCAGCG AATTTACATT TGTGTGGTGC      360

CAATCCAGAG TGAATCCGAA ACAAACTCCA TCTAGATCGC CAACCAGCAT CACGCTCGCA      420

AACGCCCCCA GAATGTACAA A ATG TTT AGG AAA CAT TTT CGG CGA AAA CCA         471
                       Met Phe Arg Lys His Phe Arg Arg Lys Pro
                         1               5                  10

GCT ACG TCG TCG TCG TTG GAG TCA ACA ATA GAA TCA GCA GAC AGC CTG         519
Ala Thr Ser Ser Ser Leu Glu Ser Thr Ile Glu Ser Ala Asp Ser Leu
                    15                  20                  25

GGA ATG TCC AAG AAG ACG GCG ACA AAA AGG CAG CGT CCG AGG CAT CGG         567
Gly Met Ser Lys Lys Thr Ala Thr Lys Arg Gln Arg Pro Arg His Arg
               30                  35                  40

GTA CCC AAA ATC GCG ACC CTG CCA TCG ACG ATC CGC GAT TGT CGA TCA         615
Val Pro Lys Ile Ala Thr Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser
          45                  50                  55

TTA AAG TCT GCC TGC AAC TTA ATT GCT TTA ATT TTA ATA CTG TTA GTC         663
Leu Lys Ser Ala Cys Asn Leu Ile Ala Leu Ile Leu Ile Leu Leu Val
     60                  65                  70

CAT AAG ATA TCC GCA GCT GGT AAC TTC GAG CTG GAA ATA TTA GAA ATC         711
His Lys Ile Ser Ala Ala Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile
 75                  80                  85                  90

TCA AAT ACC AAC AGC CAT CTA CTC AAC GGC TAT TGC TGC GGC ATG CCA         759
Ser Asn Thr Asn Ser His Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro
                95                 100                 105

GCG GAA CTT AGG GCC ACC AAG ACG ATA GGC TGC TCG CCA TGC ACG ACG         807
Ala Glu Leu Arg Ala Thr Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr
           110                 115                 120

GCA TTC CGG CTG TGC CTG AAG GAG TAC CAG ACC ACG GAG CAG GGT GCC         855
Ala Phe Arg Leu Cys Leu Lys Glu Tyr Gln Thr Thr Glu Gln Gly Ala
      125                 130                 135

AGC ATA TCC ACG GGC TGT TCG TTT GGC AAC GCC ACC ACC AAG ATA CTG         903
Ser Ile Ser Thr Gly Cys Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu
  140                 145                 150

GGT GGC TCC AGC TTT GTG CTC AGC GAT CCG GGT GTG GGA GCC ATT GTG         951
Gly Gly Ser Ser Phe Val Leu Ser Asp Pro Gly Val Gly Ala Ile Val
155                 160                 165                 170

CTG CCC TTT ACG TTT CGT TGG ACG AAG TCG TTT ACG CTG ATA CTG CAG         999
Leu Pro Phe Thr Phe Arg Trp Thr Lys Ser Phe Thr Leu Ile Leu Gln
                175                 180                 185

GCG TTG GAT ATG TAC AAC ACA TCC TAT CCA GAT GCG GAG AGG TTA ATT        1047
Ala Leu Asp Met Tyr Asn Thr Ser Tyr Pro Asp Ala Glu Arg Leu Ile
           190                 195                 200

GAG GAA ACA TCA TAC TCG GGC GTG ATA CTG CCG TCG CCG GAG TGG AAG        1095
Glu Glu Thr Ser Tyr Ser Gly Val Ile Leu Pro Ser Pro Glu Trp Lys
      205                 210                 215

ACG CTG GAC CAC ATC GGG CGG AAC GCG CGG ATC ACC TAC CGT GTC CGG        1143
Thr Leu Asp His Ile Gly Arg Asn Ala Arg Ile Thr Tyr Arg Val Arg
  220                 225                 230

GTG CAA TGC GCC GTT ACC TAC TAC AAC ACG ACC TGC ACG ACC TTC TGC        1191
Val Gln Cys Ala Val Thr Tyr Tyr Asn Thr Thr Cys Thr Thr Phe Cys
235                 240                 245                 250

CGT CCG CGG GAC GAT CAG TTC GGT CAC TAC GCC TGC GGC TCC GAG GGT        1239
Arg Pro Arg Asp Asp Gln Phe Gly His Tyr Ala Cys Gly Ser Glu Gly
```

```
                    255                 260                 265

CAG AAG CTC TGC CTG AAT GGC TGG CAG GGC GTC AAC TGC GAG GAG GCC      1287
Gln Lys Leu Cys Leu Asn Gly Trp Gln Gly Val Asn Cys Glu Glu Ala
            270                 275                 280

ATA TGC AAG GCG GGC TGC GAC CCC GTC CAC GGC AAG TGC GAT CGT CCG      1335
Ile Cys Lys Ala Gly Cys Asp Pro Val His Gly Lys Cys Asp Arg Pro
        285                 290                 295

GGG GAA TGC GAA TGC AGA CCC GGC TGG CGT GGT CCA TTG TGC AAC GAG      1383
Gly Glu Cys Glu Cys Arg Pro Gly Trp Arg Gly Pro Leu Cys Asn Glu
    300                 305                 310

TGC ATG GTC TAT CCC GGC TGC AAG CAT GGT TCC TGC AAC GGC AGC GCC      1431
Cys Met Val Tyr Pro Gly Cys Lys His Gly Ser Cys Asn Gly Ser Ala
315                 320                 325                 330

TGG AAA TGC GTG TGC GAC ACC AAC TGG GGT GGC ATA TTG TGC GAT CAA      1479
Trp Lys Cys Val Cys Asp Thr Asn Trp Gly Gly Ile Leu Cys Asp Gln
                335                 340                 345

GAT TTA AAT TTC TGC GGC ACC CAT GAA CCC TGC AAG CAC GGC GGC ACC      1527
Asp Leu Asn Phe Cys Gly Thr His Glu Pro Cys Lys His Gly Gly Thr
            350                 355                 360

TGC GAA AAT ACC GCT CCG GAC AAA TAT CGG TGC ACA TGC GCC GAG GGC      1575
Cys Glu Asn Thr Ala Pro Asp Lys Tyr Arg Cys Thr Cys Ala Glu Gly
        365                 370                 375

CTC TCG GGC GAG CAG TGC GAG ATC GTG GAG CAC CCA TGT GCC ACC AGG      1623
Leu Ser Gly Glu Gln Cys Glu Ile Val Glu His Pro Cys Ala Thr Arg
    380                 385                 390

CCA TGC CGC AAC GGC GGC ACA TGC ACA CTC AAG ACG AGT AAC CGA ACT      1671
Pro Cys Arg Asn Gly Gly Thr Cys Thr Leu Lys Thr Ser Asn Arg Thr
395                 400                 405                 410

CAA GCC CAA GTG TAT CGC ACA TCA CAT GGC AGG AGC AAC ATG GGC CGG      1719
Gln Ala Gln Val Tyr Arg Thr Ser His Gly Arg Ser Asn Met Gly Arg
                415                 420                 425

CCG GTA AGA CGC AGC AGT TCG ATG CGC AGC CTG GAT CAC CTG CGG CCG      1767
Pro Val Arg Arg Ser Ser Ser Met Arg Ser Leu Asp His Leu Arg Pro
            430                 435                 440

GAG GGG CAG GCG CTG AAT GGC AGC AGC TCC TCG GGA TTG GTG TCC CTA      1815
Glu Gly Gln Ala Leu Asn Gly Ser Ser Ser Ser Gly Leu Val Ser Leu
        445                 450                 455

GGT TCG CTG CAG CTG CAG CAG CAA CTG GCC CCC GAC TTC ACT TGC GAC      1863
Gly Ser Leu Gln Leu Gln Gln Gln Leu Ala Pro Asp Phe Thr Cys Asp
    460                 465                 470

TGC GCA GCC GGA TGG ACG GGA CCG ACA TGC GAA ATA AAT ATC GAC GAG      1911
Cys Ala Ala Gly Trp Thr Gly Pro Thr Cys Glu Ile Asn Ile Asp Glu
475                 480                 485                 490

TGC GCC GGG GGT CCC TGC GAG CAT GGT GGC ACT TGC ATC GAT CTA ATC      1959
Cys Ala Gly Gly Pro Cys Glu His Gly Gly Thr Cys Ile Asp Leu Ile
                495                 500                 505

GGT GGC TTT CGA TGT GAA TGT CCG CCG GAG TGG CAT GGC GAT GTC TGT      2007
Gly Gly Phe Arg Cys Glu Cys Pro Pro Glu Trp His Gly Asp Val Cys
            510                 515                 520

CAG GTG GAT GTG AAC GAG TGC GAG GCG CCG CAT TCC GCC GGA ATC GCT      2055
Gln Val Asp Val Asn Glu Cys Glu Ala Pro His Ser Ala Gly Ile Ala
        525                 530                 535
```

```
GCG AAC GCA TTG CTG ACC ACC ACA GCC ACC GCG ATT ATT GGT AGT AAT      2103
Ala Asn Ala Leu Leu Thr Thr Thr Ala Thr Ala Ile Ile Gly Ser Asn
    540                 545                 550

CTG AGC AGT ACT GCT CTT CTG GCC GCT CTG ACC AGT GCA GTG GCA TCC      2151
Leu Ser Ser Thr Ala Leu Leu Ala Ala Leu Thr Ser Ala Val Ala Ser
555                 560                 565                 570

ACA TCC TTG GCC ATC GGA CCC TGC ATC AAT GCC AAG GAG TGT CGC AAT      2199
Thr Ser Leu Ala Ile Gly Pro Cys Ile Asn Ala Lys Glu Cys Arg Asn
                575                 580                 585

CAG CCG GGT TCC TTT GCC TGC ATC TGC AAG GAG GGC TGG GGC GGA GTG      2247
Gln Pro Gly Ser Phe Ala Cys Ile Cys Lys Glu Gly Trp Gly Gly Val
            590                 595                 600

ACC TGT GCC GAG AAT CTA GAT GAC TGT GTG GGT CAG TGC CGG AAT GGA      2295
Thr Cys Ala Glu Asn Leu Asp Asp Cys Val Gly Gln Cys Arg Asn Gly
        605                 610                 615

GCC ACC TGC ATT GAT CTG GTC AAC GAC TAT AGG TGC GCC TGT GCC TCT      2343
Ala Thr Cys Ile Asp Leu Val Asn Asp Tyr Arg Cys Ala Cys Ala Ser
    620                 625                 630

GGA TTC ACG GGT CGC GAT TGC GAG ACG GAC ATA GAC GAG TGC GCC ACT      2391
Gly Phe Thr Gly Arg Asp Cys Glu Thr Asp Ile Asp Glu Cys Ala Thr
635                 640                 645                 650

TCC CCG TGC CGA AAC GGA GGC GAA TGT GTG GAC ATG GTG GGC AAA TTC      2439
Ser Pro Cys Arg Asn Gly Gly Glu Cys Val Asp Met Val Gly Lys Phe
                655                 660                 665

AAT TGC ATT TGC CCA CTT GGC TAC TCG GGT TCT CTG TGC GAG GAG GCC      2487
Asn Cys Ile Cys Pro Leu Gly Tyr Ser Gly Ser Leu Cys Glu Glu Ala
            670                 675                 680

AAG GAG AAC TGC ACA CCG TCG CCA TGT TTG GAG GGT CAC TGC CTC AAC      2535
Lys Glu Asn Cys Thr Pro Ser Pro Cys Leu Glu Gly His Cys Leu Asn
        685                 690                 695

ACG CCC GAA GGA TAC TAC TGC CAT TGT CCA CCG GAT CGC GCC GGA AAG      2583
Thr Pro Glu Gly Tyr Tyr Cys His Cys Pro Pro Asp Arg Ala Gly Lys
    700                 705                 710

CAC TGC GAG CAA CTG CGT CCG CTC TGC TCC CAG CCG CCC TGC AAC GAG      2631
His Cys Glu Gln Leu Arg Pro Leu Cys Ser Gln Pro Pro Cys Asn Glu
715                 720                 725                 730

GGC TGC TTC GCC AAT GTC AGC CTA GCG ACG TCA GCG ACA ACG ACG ACG      2679
Gly Cys Phe Ala Asn Val Ser Leu Ala Thr Ser Ala Thr Thr Thr Thr
                735                 740                 745

ACA ACC ACC ACA ACG GCG ACA ACG ACA AGG AAG ATG GCC AAG CCA AGC      2727
Thr Thr Thr Thr Thr Ala Thr Thr Thr Arg Lys Met Ala Lys Pro Ser
            750                 755                 760

GGA TTG CCC TGC AGC GGA CAC GGC AGC TGC GAG ATG AGC GAC GTG GGC      2775
Gly Leu Pro Cys Ser Gly His Gly Ser Cys Glu Met Ser Asp Val Gly
        765                 770                 775

ACC TTC TGC AAA TGC CAT GTG GGC CAC ACC GGC ACC TTC TGC GAG CAC      2823
Thr Phe Cys Lys Cys His Val Gly His Thr Gly Thr Phe Cys Glu His
    780                 785                 790

AAT CTC AAC GAA TGC TCG CCG AAT CCT TGT CGA AAT GGG GGA ATT TGC      2871
Asn Leu Asn Glu Cys Ser Pro Asn Pro Cys Arg Asn Gly Gly Ile Cys
795                 800                 805                 810

CTT GAC GGC GAC GGC GAT TTT ACA TGC GAG TGC ATG TCG GGC TGG ACA      2919
Leu Asp Gly Asp Gly Asp Phe Thr Cys Glu Cys Met Ser Gly Trp Thr
```

```
                    815                 820                 825

GGT AAA CGC TGC TCG GAG CGC GCT ACA GGT TGT TAT GCC GGT CAG TGC      2967
Gly Lys Arg Cys Ser Glu Arg Ala Thr Gly Cys Tyr Ala Gly Gln Cys
            830                 835                 840

CAG AAT GGT GGT ACC TGC ATG CCT GGA GCC CCG GAC AAG GCT CTG CAG      3015
Gln Asn Gly Gly Thr Cys Met Pro Gly Ala Pro Asp Lys Ala Leu Gln
        845                 850                 855

CCG CAT TGC CGC TGT GCG CCA GGT TGG ACT GGT CTG TTT TGC GCC GAG      3063
Pro His Cys Arg Cys Ala Pro Gly Trp Thr Gly Leu Phe Cys Ala Glu
    860                 865                 870

GCT ATT GAC CAG TGT CGC GGG CAG CCG TGC CAC AAT GGC GGA ACG TGC      3111
Ala Ile Asp Gln Cys Arg Gly Gln Pro Cys His Asn Gly Gly Thr Cys
875                 880                 885                 890

GAG TCG GGA GCG GGC TGG TTC CGC TGC GTC TGC GCT CAG GGA TTC TCT      3159
Glu Ser Gly Ala Gly Trp Phe Arg Cys Val Cys Ala Gln Gly Phe Ser
                895                 900                 905

GGT CCA GAC TGC CGC ATC AAT GTG AAC GAG TGC TCG CCA CAG CCT TGC      3207
Gly Pro Asp Cys Arg Ile Asn Val Asn Glu Cys Ser Pro Gln Pro Cys
            910                 915                 920

CAG GGC GGT GCC ACC TGC ATC GAC GGA ATC GGT GGA TAC AGC TGC ATC      3255
Gln Gly Gly Ala Thr Cys Ile Asp Gly Ile Gly Gly Tyr Ser Cys Ile
        925                 930                 935

TGC CCA CCA GGA AGG CAT GGA TTG CGG TGT GAA ATT TTG CTC TCC GAT      3303
Cys Pro Pro Gly Arg His Gly Leu Arg Cys Glu Ile Leu Leu Ser Asp
    940                 945                 950

CCC AAG TCC GCC TGC CAG AAC GCA AGC AAC ACT ATC TCT CCG TAT ACA      3351
Pro Lys Ser Ala Cys Gln Asn Ala Ser Asn Thr Ile Ser Pro Tyr Thr
955                 960                 965                 970

GCT CTA AAC CGA AGC CAA AAC TGG CTG GAT ATT GCT CTA ACC GGA AGA      3399
Ala Leu Asn Arg Ser Gln Asn Trp Leu Asp Ile Ala Leu Thr Gly Arg
                975                 980                 985

ACA GAA GAC GAT GAG AAC TGC AAT GCG TGT GTC TGC GAA AAC GGC ACC      3447
Thr Glu Asp Asp Glu Asn Cys Asn Ala Cys Val Cys Glu Asn Gly Thr
            990                 995                 1000

TCT CGG TGC ACG AAT CTC TGG TGT GGA TTG CCC AAT TGC TAT AAG GTG      3495
Ser Arg Cys Thr Asn Leu Trp Cys Gly Leu Pro Asn Cys Tyr Lys Val
        1005                1010                1015

GAT CCG CTC TCC AAG TCC TCG AAT CTG TCC GGT GTT TGC AAA CAG CAC      3543
Asp Pro Leu Ser Lys Ser Ser Asn Leu Ser Gly Val Cys Lys Gln His
    1020                1025                1030

GAG GTG TGC GTT CCG GCA CTG AGT GAG ACA TGC CTG TCA TCG CCT TGT      3591
Glu Val Cys Val Pro Ala Leu Ser Glu Thr Cys Leu Ser Ser Pro Cys
1035                1040                1045                1050

AAT GTT CGT GGA GAT TGC CGG GCA CTG GAA CCA TCG CGT CGG GTT GCT      3639
Asn Val Arg Gly Asp Cys Arg Ala Leu Glu Pro Ser Arg Arg Val Ala
                1055                1060                1065

CCA CCC CGA CTG CCA GCC AAA TCT AGC TGC TGG CCC AAT CAG GCC GTG      3687
Pro Pro Arg Leu Pro Ala Lys Ser Ser Cys Trp Pro Asn Gln Ala Val
            1070                1075                1080

GTC AAC GAG AAC TGC GCC CGA CTC ACC ATC CTT TTG GCC CTG GAG CGA      3735
Val Asn Glu Asn Cys Ala Arg Leu Thr Ile Leu Leu Ala Leu Glu Arg
        1085                1090                1095
```

```
GTG GGC AAG GGA GCT TCG GTG GAG GGT CTC TGC TCC CTG GTA AGG GTG      3783
Val Gly Lys Gly Ala Ser Val Glu Gly Leu Cys Ser Leu Val Arg Val
    1100                1105                1110

CTG CTG GCT GCC CAG TTG ATC AAG AAG CCG GCG AGT ACT TTT GGC CAG      3831
Leu Leu Ala Ala Gln Leu Ile Lys Lys Pro Ala Ser Thr Phe Gly Gln
1115                1120                1125                1130

GAT CCG GGA ATG CTT ATG GTG CTC TGC GAT CTC AAA ACG GGC ACC AAT      3879
Asp Pro Gly Met Leu Met Val Leu Cys Asp Leu Lys Thr Gly Thr Asn
                1135                1140                1145

GAT ACC GTT GAA CTA ACT GTG TCG TCC AGT AAA TTA AAT GAT CCC CAG      3927
Asp Thr Val Glu Leu Thr Val Ser Ser Ser Lys Leu Asn Asp Pro Gln
            1150                1155                1160

CTG CCA GTG GCG GTG GGT CTG CTG GGT GAA CTC CTG AGC TCC AGG CAG      3975
Leu Pro Val Ala Val Gly Leu Leu Gly Glu Leu Leu Ser Ser Arg Gln
        1165                1170                1175

TTG AAT GGC ATC CAG CGG CGC AAG GAA CTG GAG CTG CAG CAT GCA AAA      4023
Leu Asn Gly Ile Gln Arg Arg Lys Glu Leu Glu Leu Gln His Ala Lys
    1180                1185                1190

TTG GCT GCC CTC ACC TCC ATT GTG GAG GTC AAG TTG GAA ACG GCC CGC      4071
Leu Ala Ala Leu Thr Ser Ile Val Glu Val Lys Leu Glu Thr Ala Arg
1195                1200                1205                1210

GTG GCC GAT GGA TCG GGT CAT AGT CTG CTG ATA GGA GTG CTA TGC GGT      4119
Val Ala Asp Gly Ser Gly His Ser Leu Leu Ile Gly Val Leu Cys Gly
                1215                1220                1225

GTC TTT ATA GTC CTG GTG GGA TTC TCG GTG TTC ATC AGT CTT TAC TGG      4167
Val Phe Ile Val Leu Val Gly Phe Ser Val Phe Ile Ser Leu Tyr Trp
            1230                1235                1240

AAA CAG CGT CTG GCT TAT CGC ACC AGT TCG GGA ATG AAC TTA ACT CCC      4215
Lys Gln Arg Leu Ala Tyr Arg Thr Ser Ser Gly Met Asn Leu Thr Pro
        1245                1250                1255

TCC CTG GAT GCA CTG CGT CAC GAG GAG GAG AAG TCG AAT AAT CTG CAG      4263
Ser Leu Asp Ala Leu Arg His Glu Glu Glu Lys Ser Asn Asn Leu Gln
    1260                1265                1270

AAC GAG GAG AAT CTG CGA AGG TAT ACA AAT CCG CTG AAG GGC AGC ACC      4311
Asn Glu Glu Asn Leu Arg Arg Tyr Thr Asn Pro Leu Lys Gly Ser Thr
1275                1280                1285                1290

AGT TCC CTA AGA GCG GCC ACC GGC ATG GAA CTA AGC CTC AAT CCC GCT      4359
Ser Ser Leu Arg Ala Ala Thr Gly Met Glu Leu Ser Leu Asn Pro Ala
                1295                1300                1305

CCG GAA TTA GCC GCC TCG GCG GCG AGT AGT TCC GCC TTG CAC AGA TCG      4407
Pro Glu Leu Ala Ala Ser Ala Ala Ser Ser Ser Ala Leu His Arg Ser
            1310                1315                1320

CAG CCA CTA TTC CCG CCA TGC GAT TTC GAG CGT GAG CTG GAC TCC AGT      4455
Gln Pro Leu Phe Pro Pro Cys Asp Phe Glu Arg Glu Leu Asp Ser Ser
        1325                1330                1335

ACG GGC CTG AAG CAG GCG CAC AAG CGG AGC TCA CAG ATT CTG CTG CAC      4503
Thr Gly Leu Lys Gln Ala His Lys Arg Ser Ser Gln Ile Leu Leu His
    1340                1345                1350

AAA ACC CAA AAC TCG GAC ATG CGG AAG AAC ACT GTG GGC TCG CTG GAC      4551
Lys Thr Gln Asn Ser Asp Met Arg Lys Asn Thr Val Gly Ser Leu Asp
1355                1360                1365                1370

AGT CCG CGT AAG GAC TTT GGC AAG CGG TCG ATC AAC TGC AAG TCC ATG      4599
Ser Pro Arg Lys Asp Phe Gly Lys Arg Ser Ile Asn Cys Lys Ser Met
```

```
                1375                1380                1385

CCA CCC TCT TCG GGC GAC GAG GGC TCC GAT GTC CTT GCC ACC ACT GTG      4647
Pro Pro Ser Ser Gly Asp Glu Gly Ser Asp Val Leu Ala Thr Thr Val
            1390                1395                1400

ATG GTT TAGCCGTGAT CTCACCAACC AACCAATCAA GAAACCAACC AGCCGCCCAC      4703
Met Val

AGCCAGCTCA AAGTTCCAAT TGCCACAGCA CGGGCGCTAT TTCCAAGTGC ATTAGTAGCG   4763

TAATTAAAAC TAGGATATTG TTAAGGATAC CAAGGTAGGC CACAACGGAG TGGCTCTGTT   4823

GAAAACGTAA AGTTCTAAAA ATCCAGGTCT CTCAGACAAA GATGAGGTAC ACAAATAAAT   4883

TGGCTAGTTA ATCAAGCATG TTATGGCCAC GGGATGGGCA AATTTATTTG TATACCTGAT   4943

CTTATCTTAA TACTAAACCA GTTTTCTACT ATTTTTTTTT TGTGGATCAA GCTTAAAAGT   5003

TCAGCTAGGC AGGCGTTTTC CGCAGTGCCA TGTCGATGTG GAAGCCCAAA ATATTTAGGT   5063

TAGATAGTGT AATTTCGAAC TCTTCTCTTC GCTAAGCAAC ATCCTACACA GTGTGATATT   5123

TAGTGTAACC CAGGCGCGCA TTTACATTCA ATTAAAGACA ATGATATATA AATATAAACG   5183

AAATCAACTC CTTGGCTAGC ACAAGCTGTA TGTATATAGT TCTCATTTAG GATCGTCGCG   5243

CTCTATATTG TGTATAAGCT GTAAATACTG TAAATTAGCA GTTACCGTTA TTGTATTTTG   5303

TCTATAGTTA GATTGGTACT ATTAAACTAA GAACCAGCCG CAACGCGTTA GACTTTAAAA   5363

GTTGTTTGCA ATTGTACGCA ATAATATAGT TTTATGCTCG TAGTTAGGTA GCTGTGTAAC   5423

CGGGTAAGAT TCAAACGATT TTGTACTGTA TTATATACCT ATCTGTGTAG TAATATTTAT   5483

TTATTATATT AAATTTGATC TAGACGCAAT AAAGTAATAT CAATAAAGAT AGTAAAAGAC   5543

ATAAAAAAAA AAAAAAAA                                                 5561
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1404 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Phe Arg Lys His Phe Arg Arg Lys Pro Ala Thr Ser Ser Ser Leu
 1               5                  10                  15

Glu Ser Thr Ile Glu Ser Ala Asp Ser Leu Gly Met Ser Lys Lys Thr
            20                  25                  30

Ala Thr Lys Arg Gln Arg Pro Arg His Arg Val Pro Lys Ile Ala Thr
        35                  40                  45

Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser Leu Lys Ser Ala Cys Asn
    50                  55                  60

Leu Ile Ala Leu Ile Leu Ile Leu Leu Val His Lys Ile Ser Ala Ala
65                  70                  75                  80

Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile Ser Asn Thr Asn Ser His
                85                  90                  95
```

Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro Ala Glu Leu Arg Ala Thr
100 105 110

Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr Ala Phe Arg Leu Cys Leu
115 120 125

Lys Glu Tyr Gln Thr Thr Glu Gln Gly Ala Ser Ile Ser Thr Gly Cys
130 135 140

Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu Gly Gly Ser Ser Phe Val
145 150 155 160

Leu Ser Asp Pro Gly Val Gly Ala Ile Val Leu Pro Phe Thr Phe Arg
165 170 175

Trp Thr Lys Ser Phe Thr Leu Ile Leu Gln Ala Leu Asp Met Tyr Asn
180 185 190

Thr Ser Tyr Pro Asp Ala Glu Arg Leu Ile Glu Glu Thr Ser Tyr Ser
195 200 205

Gly Val Ile Leu Pro Ser Pro Glu Trp Lys Thr Leu Asp His Ile Gly
210 215 220

Arg Asn Ala Arg Ile Thr Tyr Arg Val Arg Val Gln Cys Ala Val Thr
225 230 235 240

Tyr Tyr Asn Thr Thr Cys Thr Thr Phe Cys Arg Pro Arg Asp Asp Gln
245 250 255

Phe Gly His Tyr Ala Cys Gly Ser Glu Gly Gln Lys Leu Cys Leu Asn
260 265 270

Gly Trp Gln Gly Val Asn Cys Glu Glu Ala Ile Cys Lys Ala Gly Cys
275 280 285

Asp Pro Val His Gly Lys Cys Asp Arg Pro Gly Glu Cys Glu Cys Arg
290 295 300

Pro Gly Trp Arg Gly Pro Leu Cys Asn Glu Cys Met Val Tyr Pro Gly
305 310 315 320

Cys Lys His Gly Ser Cys Asn Gly Ser Ala Trp Lys Cys Val Cys Asp
325 330 335

Thr Asn Trp Gly Gly Ile Leu Cys Asp Gln Asp Leu Asn Phe Cys Gly
340 345 350

Thr His Glu Pro Cys Lys His Gly Gly Thr Cys Glu Asn Thr Ala Pro
355 360 365

Asp Lys Tyr Arg Cys Thr Cys Ala Glu Gly Leu Ser Gly Glu Gln Cys
370 375 380

Glu Ile Val Glu His Pro Cys Ala Thr Arg Pro Cys Arg Asn Gly Gly
385 390 395 400

Thr Cys Thr Leu Lys Thr Ser Asn Arg Thr Gln Ala Gln Val Tyr Arg
405 410 415

Thr Ser His Gly Arg Ser Asn Met Gly Arg Pro Val Arg Arg Ser Ser
420 425 430

Ser Met Arg Ser Leu Asp His Leu Arg Pro Glu Gly Gln Ala Leu Asn
435 440 445

Gly Ser Ser Ser Ser Gly Leu Val Ser Leu Gly Ser Leu Gln Leu Gln
450 455 460

Gln Gln Leu Ala Pro Asp Phe Thr Cys Asp Cys Ala Ala Gly Trp Thr

465 470 475 480

Gly Pro Thr Cys Glu Ile Asn Ile Asp Glu Cys Ala Gly Gly Pro Cys
485 490 495

Glu His Gly Gly Thr Cys Ile Asp Leu Ile Gly Gly Phe Arg Cys Glu
500 505 510

Cys Pro Pro Glu Trp His Gly Asp Val Cys Gln Val Asp Val Asn Glu
515 520 525

Cys Glu Ala Pro His Ser Ala Gly Ile Ala Ala Asn Ala Leu Leu Thr
530 535 540

Thr Thr Ala Thr Ala Ile Ile Gly Ser Asn Leu Ser Ser Thr Ala Leu
545 550 555 560

Leu Ala Ala Leu Thr Ser Ala Val Ala Ser Thr Ser Leu Ala Ile Gly
565 570 575

Pro Cys Ile Asn Ala Lys Glu Cys Arg Asn Gln Pro Gly Ser Phe Ala
580 585 590

Cys Ile Cys Lys Glu Gly Trp Gly Gly Val Thr Cys Ala Glu Asn Leu
595 600 605

Asp Asp Cys Val Gly Gln Cys Arg Asn Gly Ala Thr Cys Ile Asp Leu
610 615 620

Val Asn Asp Tyr Arg Cys Ala Cys Ala Ser Gly Phe Thr Gly Arg Asp
625 630 635 640

Cys Glu Thr Asp Ile Asp Glu Cys Ala Thr Ser Pro Cys Arg Asn Gly
645 650 655

Gly Glu Cys Val Asp Met Val Gly Lys Phe Asn Cys Ile Cys Pro Leu
660 665 670

Gly Tyr Ser Gly Ser Leu Cys Glu Glu Ala Lys Glu Asn Cys Thr Pro
675 680 685

Ser Pro Cys Leu Glu Gly His Cys Leu Asn Thr Pro Glu Gly Tyr Tyr
690 695 700

Cys His Cys Pro Pro Asp Arg Ala Gly Lys His Cys Glu Gln Leu Arg
705 710 715 720

Pro Leu Cys Ser Gln Pro Pro Cys Asn Glu Gly Cys Phe Ala Asn Val
725 730 735

Ser Leu Ala Thr Ser Ala Thr Thr Thr Thr Thr Thr Thr Thr Thr Ala
740 745 750

Thr Thr Thr Arg Lys Met Ala Lys Pro Ser Gly Leu Pro Cys Ser Gly
755 760 765

His Gly Ser Cys Glu Met Ser Asp Val Gly Thr Phe Cys Lys Cys His
770 775 780

Val Gly His Thr Gly Thr Phe Cys Glu His Asn Leu Asn Glu Cys Ser
785 790 795 800

Pro Asn Pro Cys Arg Asn Gly Gly Ile Cys Leu Asp Gly Asp Gly Asp
805 810 815

Phe Thr Cys Glu Cys Met Ser Gly Trp Thr Gly Lys Arg Cys Ser Glu
820 825 830

Arg Ala Thr Gly Cys Tyr Ala Gly Gln Cys Gln Asn Gly Gly Thr Cys
835 840 845

Met Pro Gly Ala Pro Asp Lys Ala Leu Gln Pro His Cys Arg Cys Ala
850 855 860

Pro Gly Trp Thr Gly Leu Phe Cys Ala Glu Ala Ile Asp Gln Cys Arg
865 870 875 880

Gly Gln Pro Cys His Asn Gly Gly Thr Cys Glu Ser Gly Ala Gly Trp
885 890 895

Phe Arg Cys Val Cys Ala Gln Gly Phe Ser Gly Pro Asp Cys Arg Ile
900 905 910

Asn Val Asn Glu Cys Ser Pro Gln Pro Cys Gln Gly Gly Ala Thr Cys
915 920 925

Ile Asp Gly Ile Gly Gly Tyr Ser Cys Ile Cys Pro Pro Gly Arg His
930 935 940

Gly Leu Arg Cys Glu Ile Leu Leu Ser Asp Pro Lys Ser Ala Cys Gln
945 950 955 960

Asn Ala Ser Asn Thr Ile Ser Pro Tyr Thr Ala Leu Asn Arg Ser Gln
965 970 975

Asn Trp Leu Asp Ile Ala Leu Thr Gly Arg Thr Glu Asp Asp Glu Asn
980 985 990

Cys Asn Ala Cys Val Cys Glu Asn Gly Thr Ser Arg Cys Thr Asn Leu
995 1000 1005

Trp Cys Gly Leu Pro Asn Cys Tyr Lys Val Asp Pro Leu Ser Lys Ser
1010 1015 1020

Ser Asn Leu Ser Gly Val Cys Lys Gln His Glu Val Cys Val Pro Ala
1025 1030 1035 1040

Leu Ser Glu Thr Cys Leu Ser Ser Pro Cys Asn Val Arg Gly Asp Cys
1045 1050 1055

Arg Ala Leu Glu Pro Ser Arg Arg Val Ala Pro Pro Arg Leu Pro Ala
1060 1065 1070

Lys Ser Ser Cys Trp Pro Asn Gln Ala Val Val Asn Glu Asn Cys Ala
1075 1080 1085

Arg Leu Thr Ile Leu Leu Ala Leu Glu Arg Val Gly Lys Gly Ala Ser
1090 1095 1100

Val Glu Gly Leu Cys Ser Leu Val Arg Val Leu Leu Ala Ala Gln Leu
1105 1110 1115 1120

Ile Lys Lys Pro Ala Ser Thr Phe Gly Gln Asp Pro Gly Met Leu Met
1125 1130 1135

Val Leu Cys Asp Leu Lys Thr Gly Thr Asn Asp Thr Val Glu Leu Thr
1140 1145 1150

Val Ser Ser Ser Lys Leu Asn Asp Pro Gln Leu Pro Val Ala Val Gly
1155 1160 1165

Leu Leu Gly Glu Leu Leu Ser Ser Arg Gln Leu Asn Gly Ile Gln Arg
1170 1175 1180

Arg Lys Glu Leu Glu Leu Gln His Ala Lys Leu Ala Ala Leu Thr Ser
1185 1190 1195 1200

Ile Val Glu Val Lys Leu Glu Thr Ala Arg Val Ala Asp Gly Ser Gly
1205 1210 1215

His Ser Leu Leu Ile Gly Val Leu Cys Gly Val Phe Ile Val Leu Val

1220 1225 1230

Gly Phe Ser Val Phe Ile Ser Leu Tyr Trp Lys Gln Arg Leu Ala Tyr
1235 1240 1245

Arg Thr Ser Ser Gly Met Asn Leu Thr Pro Ser Leu Asp Ala Leu Arg
1250 1255 1260

His Glu Glu Glu Lys Ser Asn Asn Leu Gln Asn Glu Glu Asn Leu Arg
1265 1270 1275 1280

Arg Tyr Thr Asn Pro Leu Lys Gly Ser Thr Ser Ser Leu Arg Ala Ala
1285 1290 1295

Thr Gly Met Glu Leu Ser Leu Asn Pro Ala Pro Glu Leu Ala Ala Ser
1300 1305 1310

Ala Ala Ser Ser Ser Ala Leu His Arg Ser Gln Pro Leu Phe Pro Pro
1315 1320 1325

Cys Asp Phe Glu Arg Glu Leu Asp Ser Ser Thr Gly Leu Lys Gln Ala
1330 1335 1340

His Lys Arg Ser Ser Gln Ile Leu Leu His Lys Thr Gln Asn Ser Asp
1345 1350 1355 1360

Met Arg Lys Asn Thr Val Gly Ser Leu Asp Ser Pro Arg Lys Asp Phe
1365 1370 1375

Gly Lys Arg Ser Ile Asn Cys Lys Ser Met Pro Pro Ser Ser Gly Asp
1380 1385 1390

Glu Gly Ser Asp Val Leu Ala Thr Thr Val Met Val
1395 1400

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 708 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGCATTGGA TTAAATGTTT ATTAACAGCA TTCATTTGCT TCACAGTCAT CGTGCAGGTT 60

CACAGTTCCG GCAGCTTTGA GTTGCGCCTG AAGTACTTCA GCAACGATCA CGGGCGGGAC 120

AACGAGGGTC GCTGCTGCAG CGGGGAGTCG GACGGAGCGA CGGGCAAGTG CCTGGGCAGC 180

TGCAAGACGC GGTTTCGCGT CTGCCTAAAG CACTACCAGG CCACCATCGA CACCACCTCC 240

CAGTGCACCT ACGGGGACGT GATCACGCCC ATTCTCGGCG AGAACTCGGT CAATCTGACC 300

GACGCCCAGC GCTTCCAGAA CAAGGGCTTC ACGAATCCCA TCCAGTTCCC CTTCTCGTTC 360

TCATGGCCGG GTACCTTCTC GCTGATCGTC GAGGCCTGGC ATGATACGAA CAATAGCGGC 420

AATGCGCGAA CCAACAAGCT CCTCATCCAG CGACTCTTGG TGCAGCAGGT ACTGGAGGTG 480

TCCTCCGAAT GGAAGACGAA CAAGTCGGAA TCGCAGTACA CGTCGCTGGA GTACGATTTC 540

CGTGTCACCT GCGATCTCAA CTACTACGGA TCCGGCTGTG CCAAGTTCTG CCGGCCCCGC 600

GACGATTCAT TTGGACACTC GACTTGCTCG GAGACGGGCG AAATTATCTG TTTGACCGGA 660

TGGCAGGGCG ATTACTGTCA CATACCCAAA TGCGCCAAAG GCTGTGAA 708

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 236 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met His Trp Ile Lys Cys Leu Leu Thr Ala Phe Ile Cys Phe Thr Val
1 5 10 15

Ile Val Gln Val His Ser Ser Gly Ser Phe Glu Leu Arg Leu Lys Tyr
20 25 30

Phe Ser Asn Asp His Gly Arg Asp Asn Glu Gly Arg Cys Cys Ser Gly
35 40 45

Glu Ser Asp Gly Ala Thr Gly Lys Cys Leu Gly Ser Cys Lys Thr Arg
50 55 60

Phe Arg Val Cys Leu Lys His Tyr Gln Ala Thr Ile Asp Thr Thr Ser
65 70 75 80

Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro Ile Leu Gly Glu Asn Ser
85 90 95

Val Asn Leu Thr Asp Ala Gln Arg Phe Gln Asn Lys Gly Phe Thr Asn
100 105 110

Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp Pro Gly Thr Phe Ser Leu
115 120 125

Ile Val Glu Ala Trp His Asp Thr Asn Asn Ser Gly Asn Ala Arg Thr
130 135 140

Asn Lys Leu Leu Ile Gln Arg Leu Leu Val Gln Gln Val Leu Glu Val
145 150 155 160

Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu Ser Gln Tyr Thr Ser Leu
165 170 175

Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu Asn Tyr Tyr Gly Ser Gly
180 185 190

Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp Ser Phe Gly His Ser Thr
195 200 205

Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu Thr Gly Trp Gln Gly Asp
210 215 220

Tyr Cys His Ile Pro Lys Cys Ala Lys Gly Cys Glu
225 230 235

International Application No: PCT/ /

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 72, line 25-35 of the description [1]

A. IDENTIFICATION OF DEPOSIT [2]

Further deposits are identified on an additional sheet ☐ [3]

Name of depositary institution [4]

American Type Collection

Address of depositary institution (including postal code and country) [4]

12301 Parklawn Drive
Rockville, MD 20852 US

Date of deposit [5]	Accession Number [6]
December 11, 1991	68876

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet ☐

This is [illegible] photocopy; [illegible] returned [illegible] for signature [illegible] (see #18.

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [9] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☐ This sheet was received with the international application when filed (to be checked by the receiving Office)

(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau [10]

was

(Authorized Officer)

Form PCT/RO/134 (January 1981)

WHAT IS CLAIMED IS:

1. A substantially purified Serrate protein.

2. The protein of claim 1 which is a human protein.

3. The protein of claim 1 which is a D. melanogaster protein.

4. The protein of claim 3 which comprises the amino acid sequence substantially as set forth in Figure 4 from amino acid numbers 81-1404 (in SEQ ID NO:2).

5. The protein of claim 3 which is encoded by plasmid SerFL as deposited with the ATCC and assigned accession number _____.

6. A substantially purified human protein encoded by a nucleic acid hybridizable to plasmid SerFL or the Serrate sequence therein as deposited with the ATCC and assigned accession number _____.

7. A substantially purified derivative or analog of the protein of claim 2, which is able to display one or more functional activities of a human or D. melanogaster Serrate protein.

8. A substantially purified derivative or analog of the protein of claim 3, which is able to display one or more functional activities of a human or D. melanogaster Serrate protein.

9. The derivative or analog of claim 8 which is able to be bound by an antibody directed against a human or D. melanogaster Serrate protein.

10. A substantially purified fragment of the protein of claim 1, which is able to be bound by an antibody directed against a human Serrate protein.

11. A substantially purified fragment of the protein of claim 2, which is able to be bound by an antibody directed against a D. melanogaster Serrate protein.

12. A substantially purified fragment of the protein of claim 3, which is able to be bound by an antibody directed against a D. melanogaster Serrate protein.

13. A molecule comprising the fragment of claim 10.

14. A substantially purified fragment of the protein of claim 2 which is able to display one or more functional activities of a human or D. melanogaster Serrate protein.

15. A substantially purified fragment of a Serrate protein comprising the extracellular domain of the protein.

16. A substantially purified fragment of a Serrate protein comprising the intracellular domain of the protein.

17. A substantially purified fragment of a Serrate protein comprising the membrane-associated region of the protein.

18. A substantially purified fragment of a Serrate protein comprising the transmembrane domain of the protein.

19. A substantially purified fragment of a Serrate protein comprising an epidermal growth factor-homologous repeat of the protein.

20. A substantially purified fragment of a Serrate protein comprising a region homologous to a Notch protein or a Delta protein, and consisting of at least six amino acids.

21. A substantially purified fragment of a Serrate protein comprising amino acid numbers 63-73, 124-134, 149-158, 195-206, 214-219, or 250-259 as shown in Figure 4 (SEQ ID NO:2).

22. A substantially purified fragment of a Serrate protein comprising the region of the protein with the greatest homology to amino acid numbers 79-282 as shown in Figure 4 (SEQ ID NO:2).

23. A chimeric protein comprising a fragment of a Serrate protein consisting of at least 6 amino acids fused via a covalent bond to an amino acid sequence of a second protein, in which the second protein is not Serrate.

24. The chimeric protein of claim 23 in which the fragment of a Serrate protein is a fragment capable of being bound by an anti-Serrate antibody.

25. The chimeric protein of claim 24 in which the Serrate protein is a D. melanogaster protein.

26. The chimeric protein of claim 23 which is able to display one or more functional activities of a human or D. melanogaster Serrate protein.

27. The chimeric protein of claim 23 in which the fragment of a Serrate protein comprises an epidermal growth factor-homologous repeat of a Serrate protein.

28. A substantially purified fragment of a Serrate protein which (a) is capable of being bound by an anti-Serrate antibody; and (b) lacks the transmembrane and intracellular domains of the protein.

29. A substantially purified fragment of a Serrate protein which (a) is capable of being bound by an anti-Serrate antibody; and (b) lacks the extracellular domain of the protein.

30. An antibody which is capable of binding the Serrate protein of claim 1.

31. An antibody which is capable of binding the Serrate protein of claim 2.

32. The antibody of claim 1 which is monoclonal.

33. A molecule comprising a fragment of the antibody of claim 30, which fragment is capable of binding a Serrate protein.

34. An isolated nucleic acid comprising a nucleotide sequence encoding a Serrate protein.

35. The nucleic acid of claim 34 which is DNA.

36. An isolated nucleic acid comprising a nucleotide sequence complementary to the nucleotide sequence of claim 34.

37. An isolated nucleic acid comprising a nucleotide sequence encoding the Serrate protein of claim 2.

38. An isolated nucleic acid comprising a nucleotide sequence encoding the Serrate protein of claim 3.

39. An isolated nucleic acid comprising a nucleotide sequence encoding the Serrate protein of claim 4.

40. An isolated nucleic acid comprising the Serrate sequence contained in plasmid SerFL as deposited with the ATCC and assigned accession number _____.

41. An isolated human nucleic acid hybridizable to plasmid SerFL or the Serrate sequence therein as deposited with the ATCC and assigned accession number _____.

42. An isolated nucleic acid comprising a fragment of a Serrate gene consisting of at least 8 nucleotides.

43. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 10.

44. An isolated nucleic acid comprising a nucleotide sequence complementary to the nucleotide sequence of claim 43.

45. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 11.

46. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 12.

47. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 14.

48. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 20.

49. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 23.

50. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 28.

51. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 29.

52. A recombinant cell containing the nucleic acid of claim 34.

53. A recombinant cell containing the nucleic acid of claim 37.

54. A recombinant cell containing the nucleic acid of claim 38.

55. A recombinant cell containing the nucleic acid of claim 42.

56. A method of producing a Serrate protein comprising growing a recombinant cell containing the nucleic acid of claim 34 such that the encoded Serrate protein is expressed by the cell, and recovering the expressed Serrate protein.

57. A method of producing a Serrate protein comprising growing a recombinant cell containing the nucleic acid of claim 37 such that the encoded Serrate protein is expressed by the cell, and recovering the expressed Serrate protein.

58. A method of producing a Serrate protein comprising growing a recombinant cell containing the nucleic acid of claim 38 such that the encoded Serrate protein is expressed by the cell, and recovering the expressed Serrate protein.

59. A method of producing a protein comprising a fragment of a Serrate protein, which method comprises growing a recombinant cell containing the nucleic acid of claim 43 such that the encoded protein is expressed by the cell, and recovering the expressed protein.

55. The product of the process of claim 56.

56. The product of the process of claim 57.

57. The product of the process of claim 58.

58. The product of the process of claim 59.

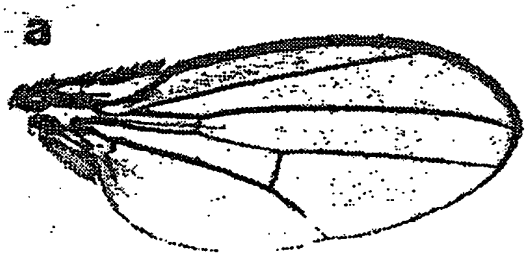


FIG. 1a

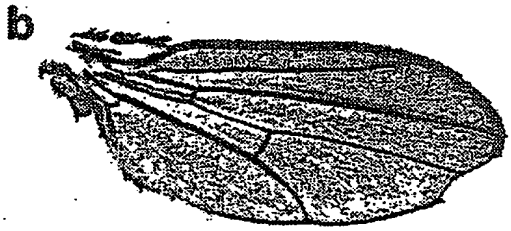


FIG. 1b

FIG. 1c

d

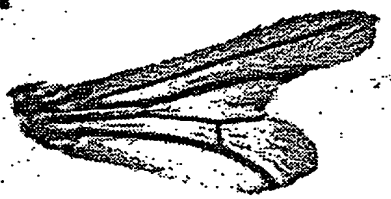


FIG. 1d

e

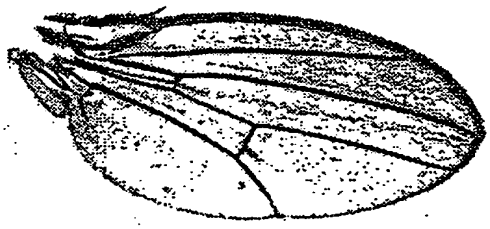


FIG. 1e

f

FIG. 1f

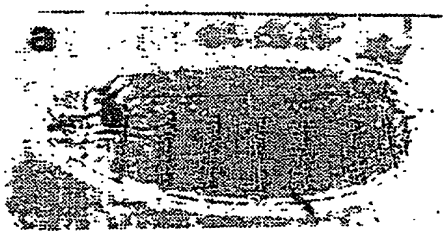


FIG.2a

FIG.2b

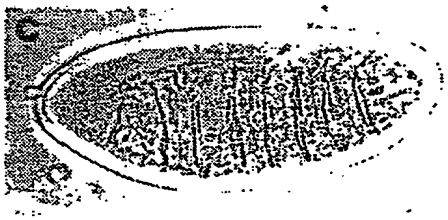


FIG.2c

FIG.2d

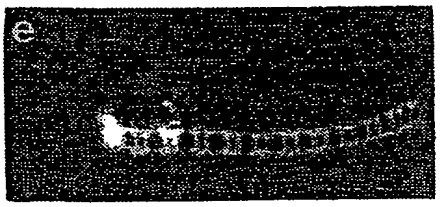


FIG.2e

FIG.2f

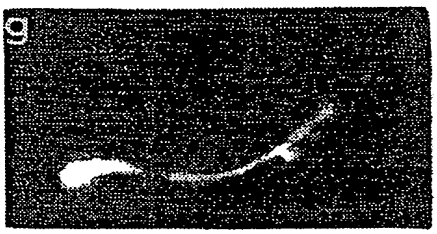


FIG.2g

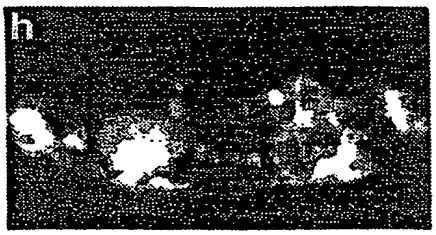


FIG.2h

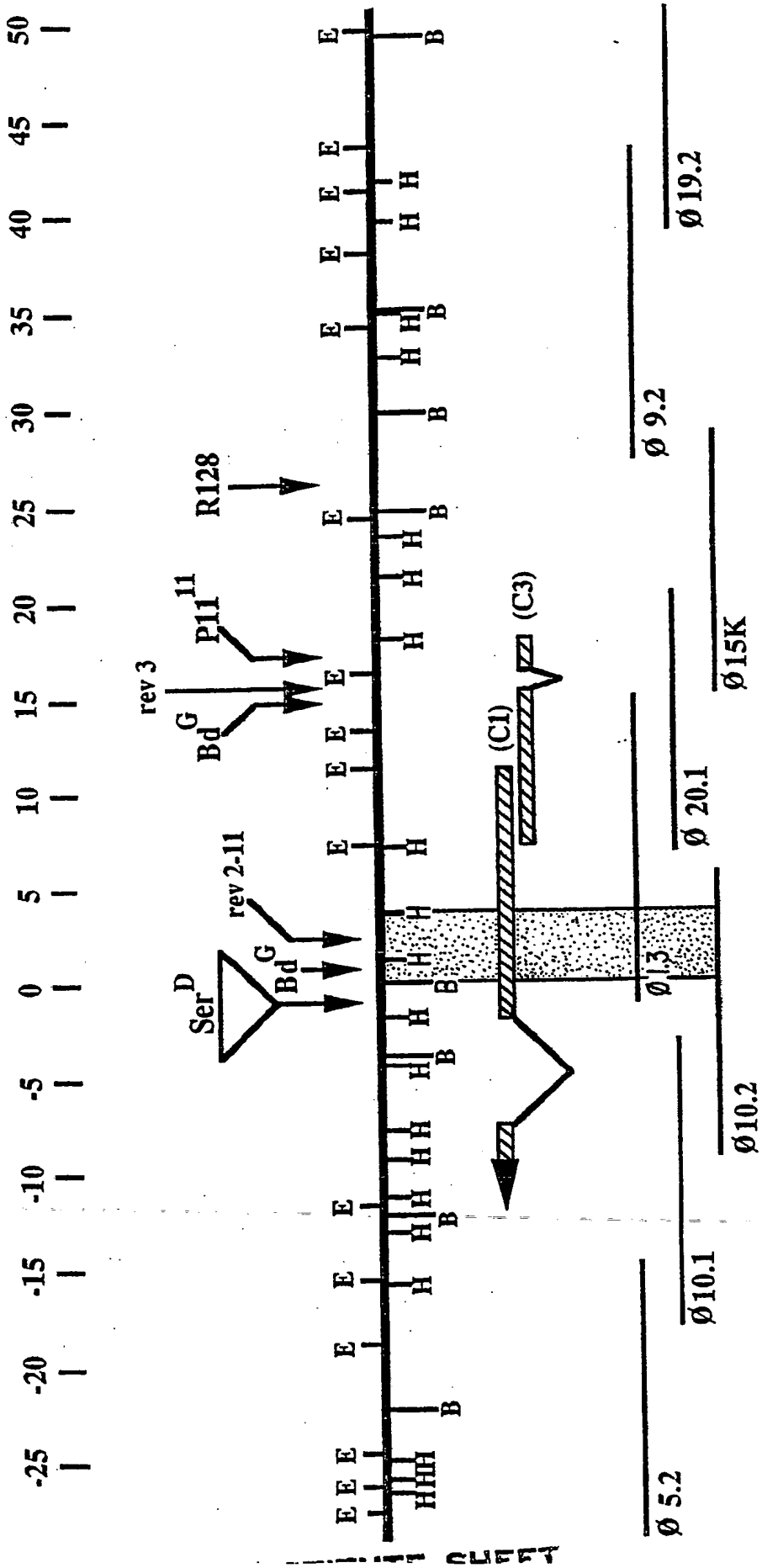

50
45
40
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
R128
P11
11
rev 3
Bd
G
rev 2-11
Ser
D
(C1)
(C3)
Ø 19.2
Ø 9.2
Ø15K
Ø 20.1
Ø13
Ø10.2
Ø10.1
Ø 5.2
E
B
H

FIG.3

```
   1 CCGAGTCGAGCGCCGTGCTTCGAGCGGTGATGAGCCCCTTTTCTGTCAACGCTAAAGATCTACAAAACATCAGCGCCTATCAAGTGGAAGTGTCAAGTGTGAACAAAACAAAAACGAGAG
 121 AAGCACATACTAAGGTCCATATAAATAATAAATAATAATTGTGTGTGATAACAACATTATCCAAACAAAACCAAACAAAACGAAGGCAAAGTGGAGAAAATGATACAGCATCCAGAGTAC
 241 GGCCGTTATTCAGCTATCCAGAGCAAGTGTAGTGTGGCAAAATAGAAACAAACAAAGGCACCAAAATCTGCATACATGGGCTAATTAAGGCTGCCCAGCGAATTTACATTTGTGTGGTGC
 361 CAATCCAGAGTGAATCCGAAACAAACTCCATCTAGATCGCCAACCAGCATCACGCTCGCAAACGCCCCCAGAATGTACAAAATGTTTAGGAAACATTTTCGGCGAAAACCAGCTACGTCG
                                                                                    MetPheArgLysHisPheArgArgLysProAlaThrSer    13

 481 TCGTCGTTGGAGTCAACAATAGAATCAGCAGACAGCCTGGGAATGTCCAAGAAGACGGCGACAAAAAGGCAGCGTCCGAGGCATCGGGTACCCAAAATCGCGACCCTGCCATCGACGATC
     SerSerLeuGluSerThrIleGluSerAlaAspSerLeuGlyMetSerLysLysThrAlaThrLysArgGlnArgProArgHisArgValProLysIleAlaThrLeuProSerThrIle    53

 601 CGCGATTGTCGATCATTAAAGTCTGCCTGCAACTTAATTGCTTTAATTTTAATACTGTTAGTCCATAAGATATCCGCAGCTGGTAACTTCGAGCTGGAAATATTAGAAATCTCAAATACC
     AspAspCysArgSerLeuLysSerAlaCysAsnLeuIleAlaLeuIleLeuIleLeuLeuValHisLysIleSerAlaAlaGlyAsnPheGluLeuGluIleLeuGluIleSerAsnThr    93

 721 AACAGCCATCTACTCAACGGCTATTGCTGCGGCATGCCAGCGGAACTTAGGGCCACCAAGACGATAGGCTGCTCGCCATGCACGACGGCATTCCGGCTGTGCCTGAAGGAGTACCAGACC
     AsnSerHisLeuLeuAsnGlyTryCysCysGlyMetProAlaGluLeuArgAlaThrLysThrIleGlyCysSerProCysThrThrAlaPheArgLeuCysLeuLysGluTyrGlnThr   133

 841 ACGGAGCAGGGTGCCAGCATATCCACGGGCTGTTCGTTTGGCAACGCCACCACCAAGATACTGGGTGGCTCCAGCTTTGTGCTCAGCGATCCGGGTGTGGGAGCCATTGTGCTGCCCTTT
     ThrGluGlnGlyAlaSerIleSerThrGlyCysSerPheGlyAsnAlaThrThrLysIleLeuGlyGlySerSerPheValLeuSerAspProGlyValGlyAlaIleValLeuProPhe   173

 961 ACGTTTCGTTGGACGAAGTCGTTTACGCTGATACTGCAGGCGTTGGATATGTACAACACATCCTATCCAGATGCGGAGAGGTTAATTGAGGAAACATCATACTCGGGCGTGATACTGCCG
     ThrPheArgTrpThrLysSerPheThrLeuIleLeuGlnAlaLeuAspMetTyrAsnThrSerTyrProAspAlaGluArgLeuIleGluGluThrSerTyrSerGlyValIleLeuPro   213
```

FIG.4a

```
1081 TCGCCGGAGTGGAAGACGCTGGACCACATCGGGCGGAACGCGCGGATCACCTACCGTGTCCGGGTGCAATGCGCCGTTACCTACTACAACACGACCTGCACGACCTTCTGCCGTCCGCGG
     SerProGluTrpLysThrLeuAspHisIleGlyArgAsnAlaArgIleThrTyrArgValArgValGlnCysAlaValThrTyrTyrAsnThrThrCysThrThrPheCysArgProArg 253

1201 GACGATCAGTTCGGTCACTACGCCTGCGGCTCCGAGGGTCAGAAGCTCTGCCTGAATGGCTGGCAGGGCGTCAACTGCGAGGAGGCCATATGCAAGGGCGGGCTGCGACCCCGTCCACGGC
     AspAspGlnPheGlyHisTyrAlaCysGlySerGluGlyGlnLysLeuCysLeuAsnGlyTrpGlnGlyValAsnCysGluGluAlaIleCysLysAlaGlyCysAspProValHisGly 293

1321 AAGTGCGATCGTCCGGGGGAATGCGAATGCAGACCCGGCTGGCGTGGTCCATTGTGCAAGGAGTGCATGGTCTATCCCGGCTGCAAGCATGGTTCCTGCAACGGCAGCGCCTGGAAATGC
     LysCysAspArgProGlyGluCysGluCysArgProGlyTrpArgGlyProLeuCysAsnGluCysMetValTyrProGlyCysLysHisGlySerCysAsnGlySerAlaTrpLysCys 333

1441 GTGTGCGACACCAACTGGGGTGGCATATTGTGCGATCAAGATTTAAATTTCTGCGGCACCCATGAACCCTGCAAGCACGGCGGCACCTGCGAAAATACCGCTCCGGACAAATATCGGTGC
     ValCysAsnThrAsnTrpGlyGlyIleLeuCysAspGlnAspLeuAsnPheCysGlyThrHisGluProCysLysHisGlyGlyThrCysGluAsnThrAlaProAspLysTyrArgCys 373

1561 ACATGCGCCGAGGGCCTCTCGGGCAGCAGTGCGAGATCGTGGAGCACCCATGTCGCCAACGGCGGCACATGCACACTCAAGACGAGTAACCGAACTCAAGCCCAA
     ThrCysAlaGluGlyLeuSerGlyGluGlnCysGluIleValGluHisProCysAlaThrArgProCysArgAsnGlyGlyThrCysThrLeuLysThrSerArgAsnArgThrGlnAlaGln 413

1681 GTGTATCGCACATCACATGGCAGGAGCAACATGGGCCGGCCGGTAAGACGCAGTTCGATGCGCAGCCTGGATCACCTGCGGCCGGAGGGGCAGGCGCTGAATGGCAGCAGCTCCTCG
     ValTyrArgThrSerHisGlyArgSerAsnMetGlyArgProValArgArgSerSerMetArgSerLeuAspHisLeuArgProGluGlyGlnAlaLeuAsnGlySerSerSerSer 453
```

FIG.4b

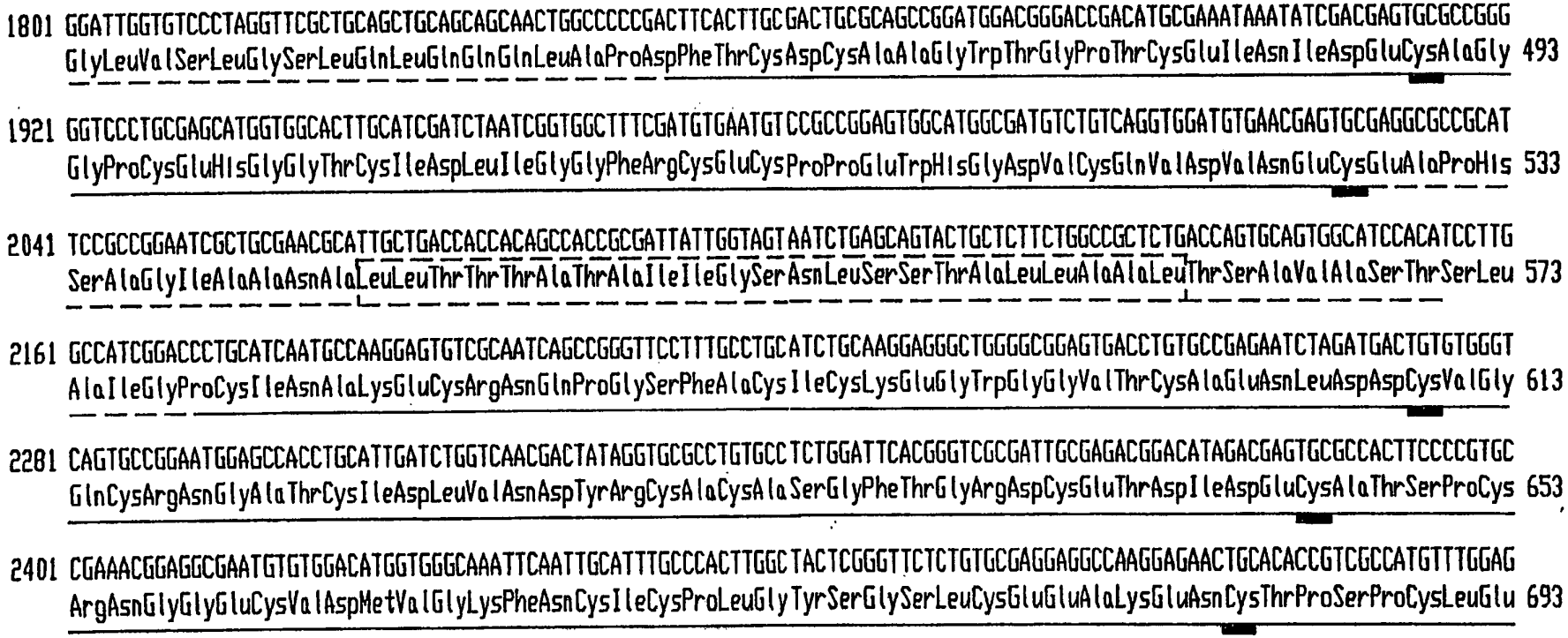

1801 GGATTGGTGTCCCTAGGTTCGCTGCAGCTGCAGCAGCAACTGGCCCCCGACTTCACTTGCGACTGCGCAGCCGGATGGACGGGACCGACATGCGAAATAAATATCGACGAGTGCGCCGGG
GlyLeuValSerLeuGlySerLeuGlnLeuGlnGlnGlnLeuAlaProAspPheThrCysAspCysAlaAlaGlyTrpThrGlyProThrCysGluIleAsnIleAspGluCysAlaGly 493
1921 GGTCCCTGCGAGCATGGTGGCACTTGCATCGATCTAATCGGTGGCTTTCGATGTGAATGTCCGCCGGAGTGGCATGGCGATGTCTGTCAGGTGGATGTGAACGAGTGCGAGGCGCCGCAT
GlyProCysGluHisGlyGlyThrCysIleAspLeuIleGlyGlyPheArgCysGluCysProProGluTrpHisGlyAspValCysGlnValAspValAsnGluCysGluAlaProHis 533
2041 TCCGCCGGAATCGCTGCGAACGCATTGCTGACCACCACAGCCACCGCGATTATTGGTAGTAATCTGAGCAGTACTGCTCTTCTGGCCGCTCTGACCAGTGCAGTGGCATCCACATCCTTG
SerAlaGlyIleAlaAlaAsnAlaLeuLeuThrThrThrAlaThrAlaIleIleGlySerAsnLeuSerSerThrAlaLeuLeuAlaAlaLeuThrSerAlaValAlaSerThrSerLeu 573
2161 GCCATCGGACCCTGCATCAATGCCAAGGAGTGTCGCAATCAGCCGGGTTCCTTTGCCTGCATCTGCAAGGAGGGCTGGGGCGGAGTGACCTGTGCCGAGAATCTAGATGACTGTGTGGGT
AlaIleGlyProCysIleAsnAlaLysGluCysArgAsnGlnProGlySerPheAlaCysIleCysLysGluGlyTrpGlyGlyValThrCysAlaGluAsnLeuAspAspCysValGly 613
2281 CAGTGCCGGAATGGAGCCACCTGCATTGATCTGGTCAACGACTATAGGTGCGCCTGTGCCTCTGGATTCACGGGTCGCGATTGCGAGACGGACATAGACGAGTGCGCCACTTCCCCGTGC
GlnCysArgAsnGlyAlaThrCysIleAspLeuValAsnAspTyrArgCysAlaCysAlaSerGlyPheThrGlyArgAspCysGluThrAspIleAspGluCysAlaThrSerProCys 653
2401 CGAAACGGAGGCGAATGTGTGGACATGGTGGGCAAATTCAATTGCATTTGCCCACTTGGCTACTCGGGTTCTCTGTGCGAGGAGGCCAAGGAGAACTGCACACCGTCGCCATGTTTGGAG
ArgAsnGlyGlyGluCysValAspMetValGlyLysPheAsnCysIleCysProLeuGlyTyrSerGlySerLeuCysGluGluAlaLysGluAsnCysThrProSerProCysLeuGlu 693

FIG.4c

2521 GGTCACTGCCTCAACACGCCCGAAGGATACTACTGCCATTGTCCACCGGATCGCGCCGGAAAGCACTGCGAGCAACTGCGTCCGCTCTGCTCCCAGCCGCCCTGCAACGAGGGCTGCTTC
GlyHisCysLeuAsnThrProGluGlyTyrTyrCysHisCysProProAspArgAlaGlyLysHisCysGluGlnLeuArgProLeuCysSerGlnProProCysAsnGluGlyCysPhe 733

2641 GCCAATGTCAGCCTAGCGACGTCAGCGACAACGACGACGACAACCACCACAACGGCGACAACGACAAGGAAGATGGCCAAGCCAAGCGGATTGCCCTGCAGCGGACACGGCAGCTGCGAG
AlaAsnValSerLeuAlaThrSerAlaThrThrThrThrThrThrThrThrAlaThrThrThrArgLysMetAlaLysProSerGlyLeuProCysSerGlyHisGlySerCysGlu 773

2761 ATGAGCGACGTGGGCACCTTCTGCAAATGCCATGTGGGCCACACCGGCACCTTCTGCGAGCACAATCTCAACGAATGCTCGCCGAATCCTTGTCGAAATGGGGGAATTTGCCTTGACGGC
MetSerAspValGlyThrPheCysLysCysHisValGlyHisThrGlyThrPheCysGluHisAsnLeuAsnGluCysSerProAsnProCysArgAsnGlyGlyIleCysLeuAspGly 813

2881 GACGGCGATTTTACATGCGAGTGCATGTCGGGCTGGACAGGTAAACGCTGCTCGGAGCGCGCTACAGGTTGTTATGCCGGTCAGTGCCAGAATGGTGGTACCTGCATGCCTGGAGCCCCG
AspGlyAspPheThrCysGluCysMetSerGlyTrpThrGlyLysArgCysSerGluArgAlaThrGlyCysTyrAlaGlyGlnCysGlnAsnGlyGlyThrCysMetProGlyAlaPro 853

3001 GACAAGGCTCTGCAGCCGCATTGCCGCTGTGCGCCAGGTTGGACTGGTCTGTTTTGCGCCGAGGCTATTGACCAGTGTCGCGGGCAGCCGTGCCACAATGGCGGAACGTGCGAGTCGGGA
AspLysAlaLeuGlnProHisCysArgCysAlaProGlyTrpThrGlyLeuPheCysAlaGluAlaIleAspGlnCysArgGlyGlnProCysHisAsnGlyGlyThrCysGluSerGly 893

3121 GCGGGCTGGTTCCGCTGCGTCTGCGCTCAGGGATTCTCTGGTCCAGACTGCCGCATCAATGTGAACGAGTGCTCGCCACAGCCTTGCCAGGGCGGTGCCACCTGCATCGACGGAATCGGT
AlaGlyTrpPheArgCysValCysAlaGlnGlyPheSerGlyProAspCysArgIleAsnValAsnGluCysSerProGlnProCysGlnGlyGlyAlaThrCysIleAspGlyIleGly 933

3241 GGATACAGCTGCATCTGCCCACCAGGAAGGCATGGATTGCGGTGTGAAATTTTGCTCTCCGATCCCAAGTCCGCCTGCCAGAACGCAAGCAACACTATCTCTCCGTATACAGCTCTAAAC
GlyTyrSerCysIleCysProProGlyArgHisGlyLeuArgCysGluIleLeuLeuSerAspProLysSerAlaCysGlnAsnAlaSerAsnThrIleSerProTyrThrAlaLeuAsn 973

FIG.4d

```
3361 CGAAGCCAAAACTGGCTGGATATTGCTCTAACCGGAAGAACAGAAGACGATGAGAACTGCAATGCGTGTGTCTGCGAAAACGGCACCTCTCGGTGCACGAATCTCTGGTGTGGATTGCCC
     ArgSerGlnAsnTrpLeuAspIleAlaLeuThrGlyArgThrGluAspAspGluAsnCysAsnAlaCysValCysGluAsnGlyThrSerArgCysThrAsnLeuTrpCysGlyLeuPro 1013

3481 AATTGCTATAAGGTGGATCCGCTCTCCAAGTCCTCGAATCTGTCCGGTGTTTGCAAACAGCACGAGGTGTGCGTTCCGGCACTGAGTGAGACATGCCTGTCATCGCCTTGTAATGTTCGT
     AsnCysTyrLysValAspProLeuSerLysSerSerAsnLeuSerGlyValCysLysGlnHisGluValCysValProAlaLeuSerGluThrCysLeuSerSerProCysAsnValArg 1053

3601 GGAGATTGCCGGGCACTGGAACCATCGCGTCGGGTTGCTCCACCCCGACTGCCAGCCAAATCTAGCTGCTGGCCCAATCAGGCCGTGGTCAACGAGAACTGCGCCCGACTCACCATCCTT
     GlyAspCysArgAlaLeuGluProSerArgArgValAlaProProArgLeuProAlaLysSerSerCysTrpProAsnGlnAlaValValAsnGluAsnCysAlaArgLeuThrIleLeu 1093

3721 TTGGCCCTGGAGCGAGTGGGCAAGGGAGCTTCGGTGGAGGGTCTCTGCTCCCTGGTAAGGGTGCTGCTGGCTGCCCAGTTGATCAAGAAGCCGGCGAGTACTTTTGGCCAGGATCCGGGA
     LeuAlaLeuGluArgValGlyLysGlyAlaSerValGluGlyLeuCysSerLeuValArgValLeuLeuAlaAlaGlnLeuIleLysLysProAlaSerThrPheGlyGlnAspProGly 1133

3841 ATGCTTATGGTGCTCTGCGATCTCAAAACGGGCACCAATGATACCGTTGAACTAACTGTGTCGTCCAGTAAATTAAATGATCCCCAGCTGCCAGTGGCGGTGGGTCTGCTGGGTGAACTC
     MetLeuMetValLeuCysAspLeuLysThrGlyThrAsnAspThrValGluLeuThrValSerSerSerLysLeuAsnAspProGlnLeuProValAlaValGlyLeuLeuGlyGluLeu 1173

3961 CTGAGCTCCAGGCAGTTGAATGGCATCCAGCGGCGCAAGGAACTGGAGCTGCAGCATGCAAAATTGGCTGCCCTCACCTCCATTGTGGAGGTCAAGTTGGAAACGGCCCGCGTGGCCGAT
     LeuSerSerArgGlnLeuAsnGlyIleGlnArgArgLysGluLeuGluLeuGlnHisAlaLysLeuAlaAlaLeuThrSerIleValGluValLysLeuGluThrAlaArgValAlaAsp 1213
```

FIG.4e

```
4081 GGATCGGGTCATAGTCTGCTGATAGGAGTGCTATGCGGTGTCTTTATAGTCCTGGTGGGATTCTCGGTGTTCATCAGTCTTTACTGGAAACAGCGTCTGGCTTATCGCACCAGTTCGGGA
     GlySerGlyHisSerLeuLeuIleGlyValLeuCysGlyValPheIleValLeuValGlyPheSerValPheIleSerLeuTyrTrpLysGlnArgLeuAlaTyrArgThrSerSerGly 1253

4201 ATGAACTTAACTCCCTCCCTGGATGCACTGCGTCACGAGGAGGAGAAGTCGAATAATCTGCAGAACGAGGAGAATCTGCGAAGGTATACAAATCCGCTGAAGGGCAGCACCAGTTCCCTA
     MetAsnLeuThrProSerLeuAspAlaLeuArgHisGluGluGluLysSerAsnAsnLeuGlnAsnGluGluAsnLeuArgArgTyrThrAsnProLeuLysGlySerThrSerSerLeu 1293

4321 AGAGCGGCCACCGGCATGGAACTAAGCCTCAATCCCGCTCCGGAATTAGCCGCCTCGGCGGCGAGTAGTTCCGCCTTGCACAGATCGCAGCCACTATTCCCGCCATGCGATTTCGAGCGT
     ArgAlaAlaThrGlyMetGluLeuSerLeuAsnProAlaProGluLeuAlaAlaSerAlaAlaSerSerSerAlaLeuHisArgSerGlnProLeuPheProProCysAspPheGluArg 1333

4441 GAGCTGGACTCCAGTACGGGCCTGAAGCAGGCGCACAAGCGGAGCTCACAGATTCTGCTGCACAAAACCCAAAACTCGGACATGCGGAAGAACACTGTGGGCTCGCTGGACAGTCCGCGT
     GluLeuAspSerSerThrGlyLeuLysGlnAlaHisLysArgSerSerGlnIleLeuLeuHisLysThrGlnAsnSerAspMetArgLysAsnThrValGlySerLeuAspSerProArg 1373

4561 AAGGACTTTGGCAAGCGGTCGATCAACTGCAAGTCCATGCCACCCTCTTCGGGCGACGAGGGCTCCGATGTCCTTGCCACCACTGTGATGGTTTAGCCGTGATCTCACCAACCAACCAAT
     LysAspPheGlyLysArgSerIleAsnCysLysSerMetProProSerSerGlyAspGluGlySerAspValLeuAlaThrThrValMetVal                          1404

4681 CAAGAAACCAACCAGCCGCCCACAGCCAGCTCAAAGTTCCAATTGCCACAGCACGGGCGCTATTTCCAAGTGCATTAGTAGCGTAATTAAAACTAGGATATTGTTAAGGATACCAAGGTA
4801 GGCCACAACGGAGTGGCTCTGTTGAAAACGTAAAGTTCTAAAAATCCAGGTCTCTCAGACAAAGATGAGGTACACAAATAAATTGGCTAGTTAATCAAGCATGTTATGGCCACGGGATGG
4921 GCAAATTTATTTGTATACCTGATCTTATCTTAATACTAAACCAGTTTTCTACTATTTTTTTTTTGTGGATCAAGCTTAAAAGTTCAGCTAGGCAGGCGTTTTCCGCAGTGCCATGTCGAT
5041 GTGGAAGCCCAAAATATTTAGGTTAGATAGTGTAATTTCGAACTCTTCTCTTCGCTAAGCAACATCCTACACAGTGTGATATTTAGTGTAACCCAGGCGCGCATTTACATTCAATTAAAG
5161 ACAATGATATATAAATATAAACGAAATCAACTCCTTGGCTAGCACAAGCTGTATGTATATAGTTCTCATTTAGGATCGTCGCGCTCTATATTGTGTATAAGCTGTAAATACTGTAAATTA
5281 GCAGTTACCGTTATTGTATTTTGTCTATAGTTAGATTGGTACTATTAAACTAAGAACCAGCCGCAACGCGTTAGACTTTAAAAGTTGTTTGCAATTGTACGCAATAATATAGTTTTATGC
5401 TCGTAGTTAGGTAGCTGTGTAACCGGGTAAGATTCAAACGATTTTGTACTGTATTATATACCTATCTGTGTAGTAATATTTATTTATTATATTAAATTTGATCTAGACGCAATAAAGTAA
5521 TATCAATAAAGATAGTAAAAGACATAAAAAAAAAAAAAAA
```

FIG.4f

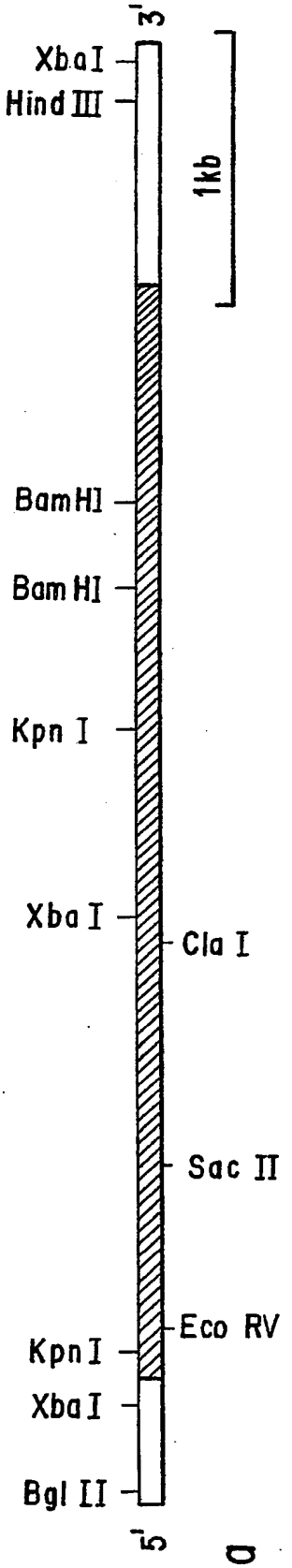


FIG.5a

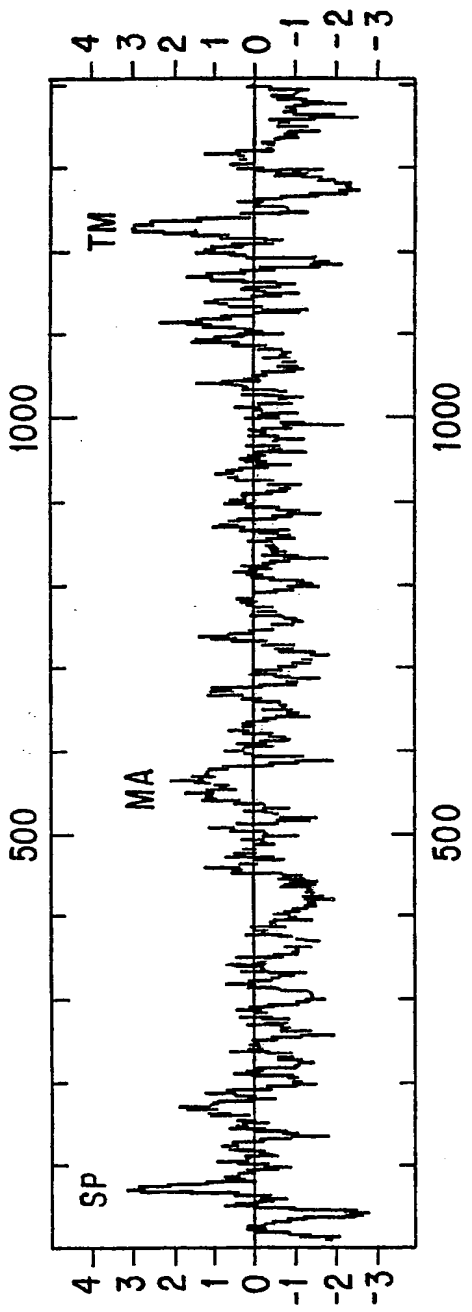


FIG.5b

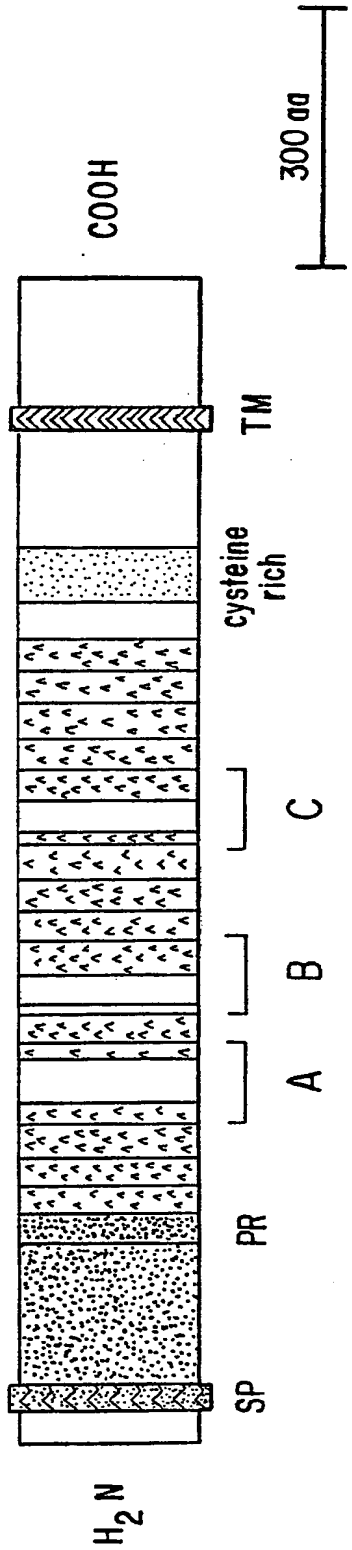


FIG.5c

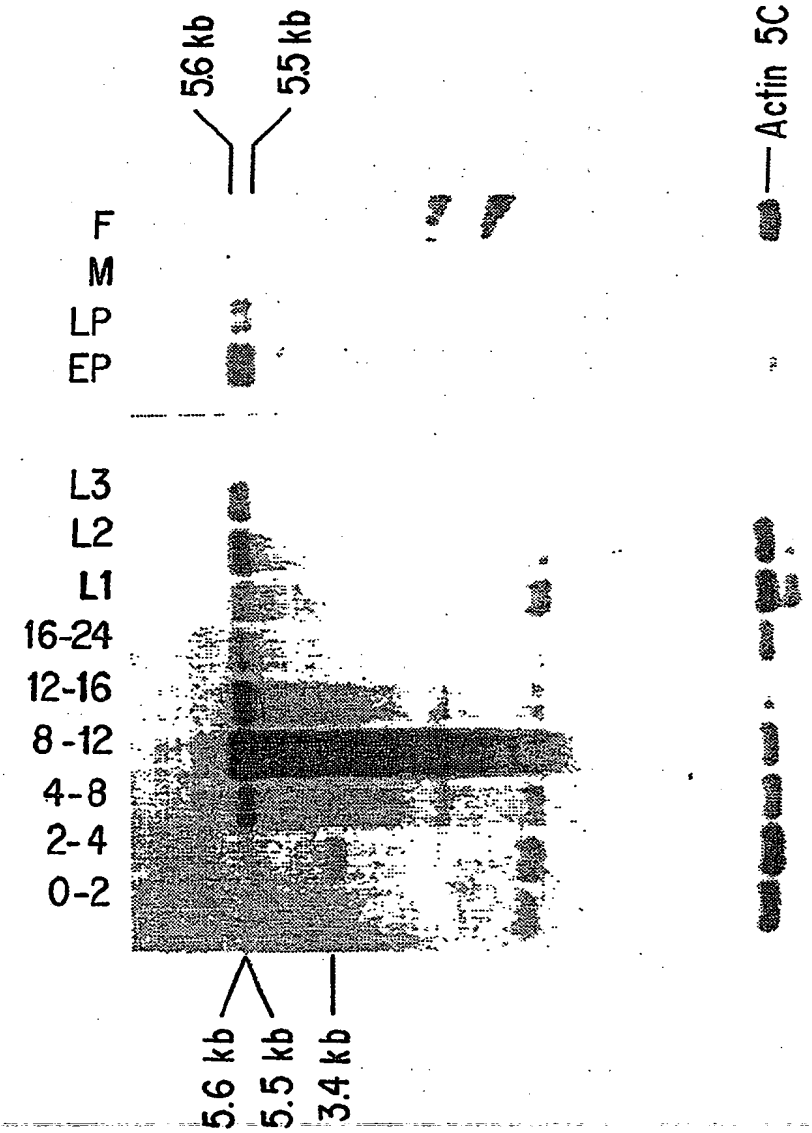

5.6 kb
5.5 kb
Actin 5C
F
M
LP
EP
L3
L2
L1
16-24
12-16
8-12
4-8
2-4
0-2
5.6 kb
5.5 kb
3.4 kb

FIG.6

FIG. 7a

FIG. 7b

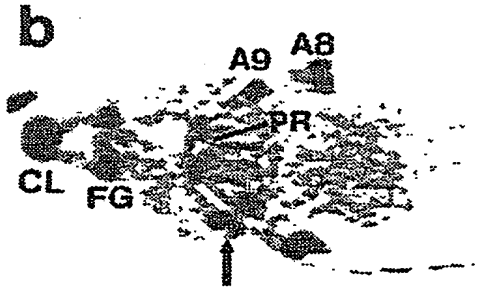


FIG. 7c

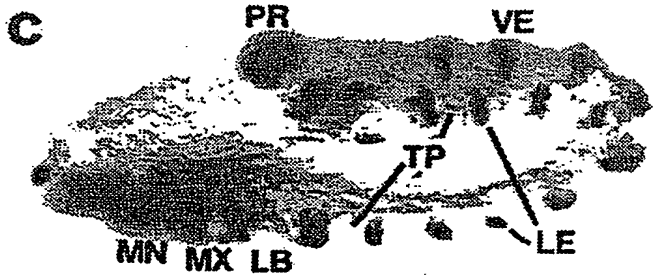


FIG. 7d

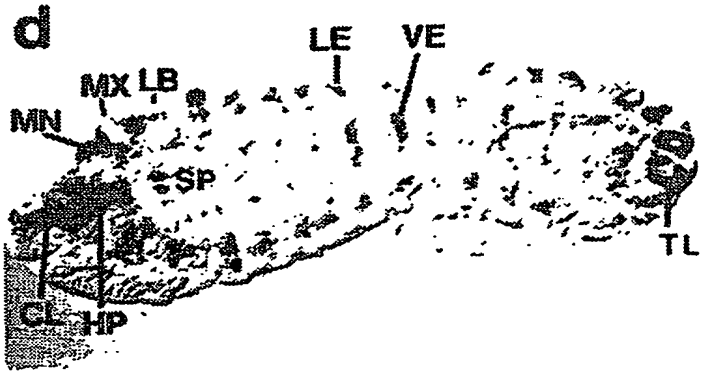


FIG.7e

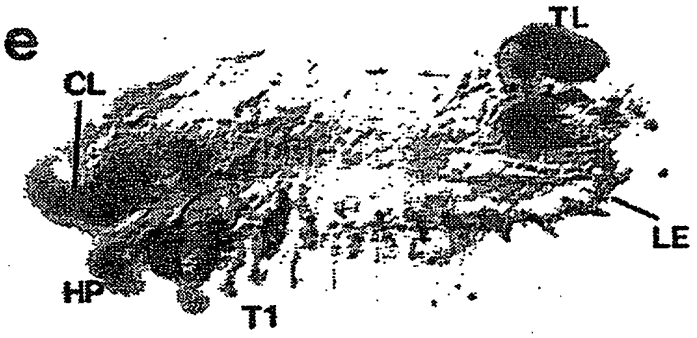


FIG.7f

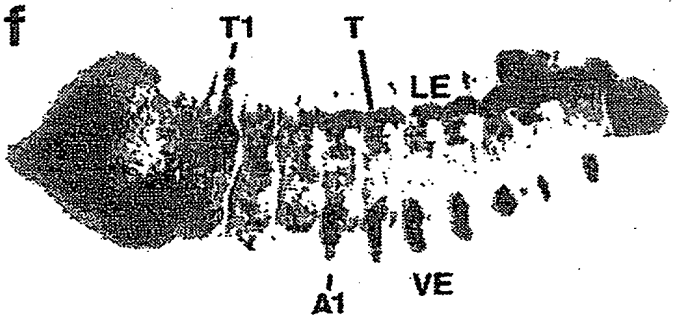


FIG.7g

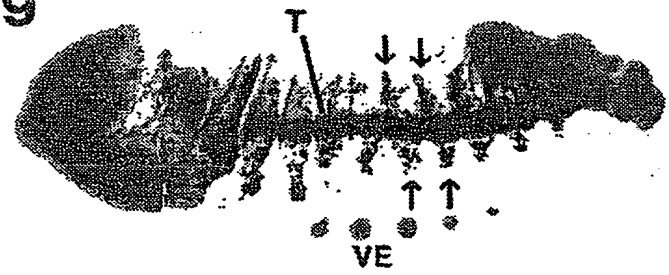


FIG.7h

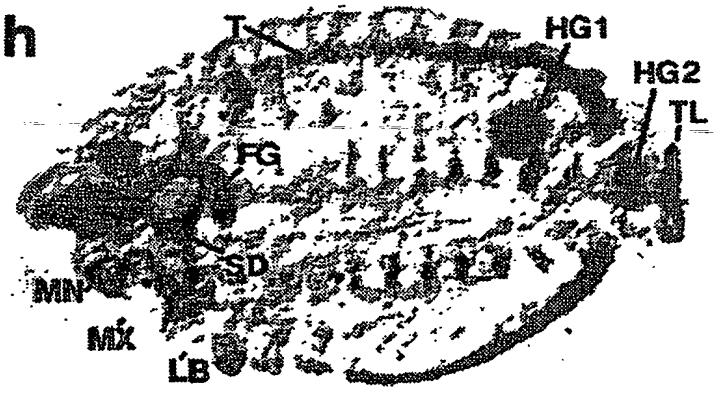


FIG.7i

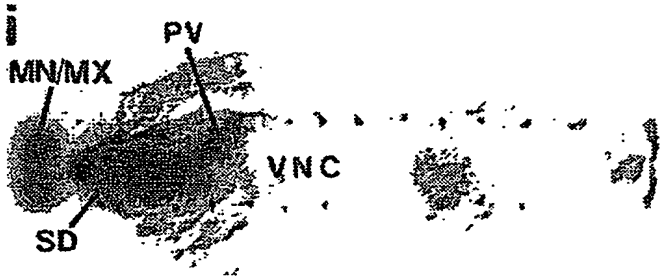


FIG.7j

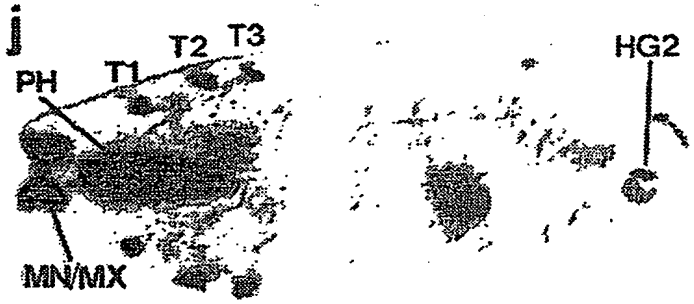


FIG.7k

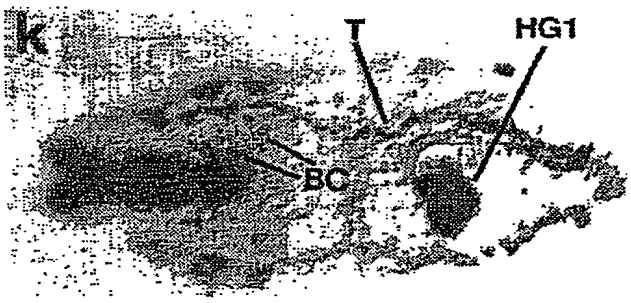
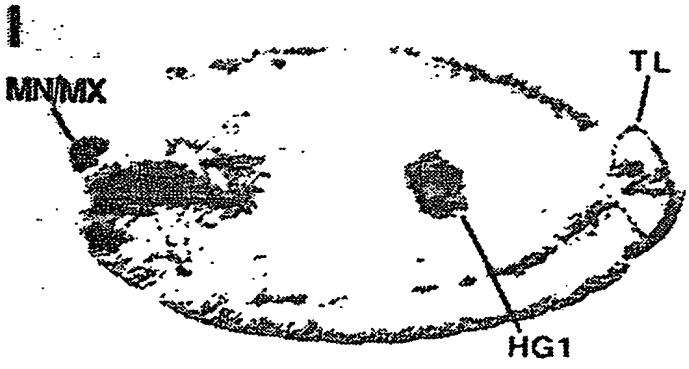


FIG.7l



SER SLKSACNLIALILILLVHKISAAGNFELEILEISNTNSHLLNGYCCGMPAELRATKTIGCSP
DELTA MHWIKCLLTAFICFTVIVQVHSSGSFELRLKYFSNDHGRDNEGRCCSGESDGATGKCLG--S
FLE1
FLE2
SER CTTAFRLCLKEYQTTEQGASISTGCSFGNATTKILGGSSFVLSDPG--------VGAIVLPFT
DELTA CKTRFRVCLKHYQATIDTTS---QCTYGDVITPILGENSVNLTDAQRFQNKGFTNPIQFPFS
FLE1R
FLE2R
FLE3
SER FRWTKSFTLILQA-LDMYNTSYPDAERLIEETSYSGVIL-PSPEWKTLDHIGRNARITYRVR
DELTA FSWPGTFSLIVEAWHDTNNSGNARTNKLLIQRLLVQQVLEVSSEWKTNKSESQYTSLEYDFR
FLE3R
FLE4
SER VQCAVTYYNTTCTTFCRPRDDQFGHYACGSEGQKLCLNGWQGVNCEEAICKAGCD
DELTA VTCDLNYYGSGCAKFCRPRDDSFGHSTCSETGEIICLTGWQGDYCHIPKCAKGCE
FLE4R

FIG.8

SERRATE TCATTAAAGTCTGCCTGCAACTTAATTGCTTTAATTTTAATACTGTTAGTCCATAAG

DELTA ATGCATTGGATTAAATGTTTATTAACAGCATTCATTTGCTTCACAGTCATCGTGCAG

SERRATE ATATCCGCAGCTGGTAACTTCGAGCTGGAAATATTAGAAATCTCAAATACCAACAGC

DELTA GTTCACAGTTCCGGCAGCTTTGAGTTGCGCCTGAAGTACTTCAGCAACGATCACGGG

SERRATE CATCTACTCAACGGCTATTGCTGCGGCATGCCAGCGGAACTTAGGGCCACCAAGACG

DELTA CGGGACAACGAGGGTCGCTGCTGCAGCGGGGAGTCGGACGGAGCGACGGGCAAGTGC

FLE1

SERRATE ATAGGCTGCTCGCCATGCACGACGGCATTCCGGCTGTGCCTGAAGGAGTACCAGACC

FLE1R

DELTA CTGGGC------AGCTGCAAGACGCGGTTTCGCGTCTGCCTAAAGCACTACCAGGCC

FIG.9A

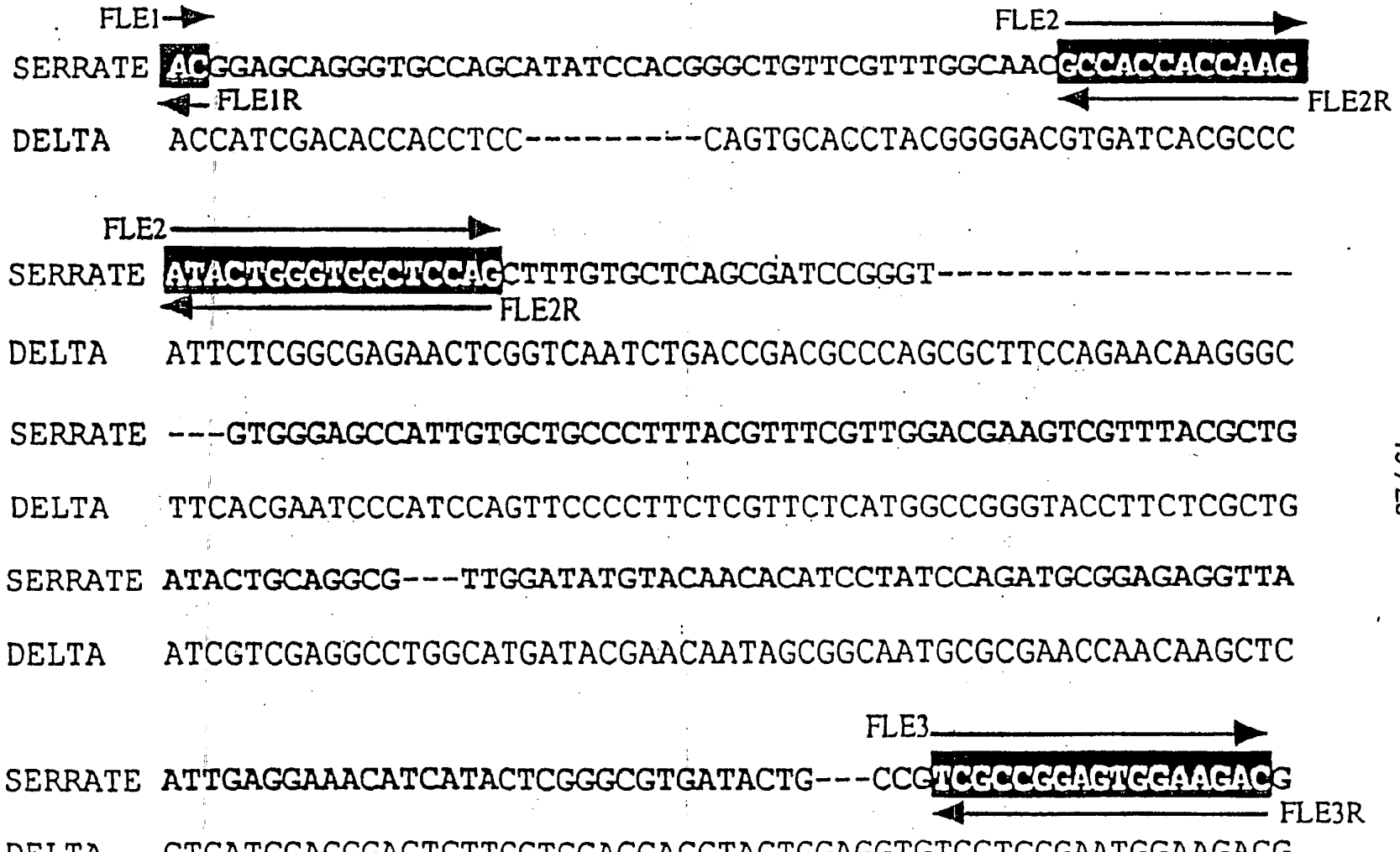

FLE1
FLE2
SERRATE ACGGAGCAGGGTGCCAGCATATCCACGGGCTGTTCGTTTGGCAACGCCACCACCAAG
FLE1R
FLE2R
DELTA ACCATCGACACCACCTCC---------CAGTGCACCTACGGGGACGTGATCACGCCC
FLE2
SERRATE ATACTGGGTGGCTCCAGCTTTGTGCTCAGCGATCCGGGT----------
FLE2R
DELTA ATTCTCGGCGAGAACTCGGTCAATCTGACCGACGCCCAGCGCTTCCAGAACAAGGGC
SERRATE ---GTGGGAGCCATTGTGCTGCCCTTTACGTTTCGTTGGACGAAGTCGTTTACGCTG
DELTA TTCACGAATCCCATCCAGTTCCCCTTCTCGTTCTCATGGCCGGGTACCTTCTCGCTG
SERRATE ATACTGCAGGCG---TTGGATATGTACAACACATCCTATCCAGATGCGGAGAGGTTA
DELTA ATCGTCGAGGCCTGGCATGATACGAACAATAGCGGCAATGCGCGAACCAACAAGCTC
FLE3
SERRATE ATTGAGGAAACATCATACTCGGGCGTGATACTG---CCGTCGCCCGGAGTGGAAGACG
FLE3R
DELTA CTCATCCAGCGACTCTTGGTGCAGCAGGTACTGGAGGTGTCCTCCGAATGGAAGACG

FIG.9B

SERRATE CTGGACCACATCGGGCGGAACGCGCGGATCACCTACCGTGTCCGGGTGCAATGCGCC

DELTA AACAAGTCGGAATCGCAGTACACGTCGCTGGAGTACGATTTCCGTGTCACCTGCGAT

FLEA

SERRATE GTTACCTACTACAACACGACCTGCACGACCTTCTGCCGTCCGCGGGACGATCAGTTC

FLEAR

DELTA CTCAACTACTACGGATCCGGCTGTGCCAAGTTCTGCCGGCCCCGCGACGATTCATTT

FLEA

SERRATE GGTCACTACGCCTGCGGCTCCGAGGGTCAGAAGCTCTGCCTGAATGGCTGGCAGGGC

FLEAR

DELTA GGACACTCGACTTGCTCGGAGACGGGCGAAATTATCTGTTTGACCGGATGGCAGGGC

SERRATE GTCAACTGCGAGGAGGCCATATGCAAGGCGGGCTGCGAC

DELTA GATTACTGTCACATACCCAAATGCGCCAAAGGCTGTGAA

FIG. 9C

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09240

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C07K 13/00; C12P 21/06
US CL : 530/387; 435/69.1

II. FIELDS SEARCHED

Minimum Documentation Searched[4]

Classification System	Classification Symbols
U.S.	530/387; 435/69.1

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched[5]

APS, DIALOG search terms: serrate
SEQUENCE SEARCH search terms: sequence of figure 4 (protein only)

III. DOCUMENTS CONSIDERED TO BE RELEVANT[14]

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]
Y	Journal of Biological Chemistry, Volume 265, No.35, Issued 15 December 1990, Y.-B. Xie et al., "Extracellular domain of lutropin/choriogonadotropin receptor expressed in transfected cells binds choriogonadotrophin with high affinity", pages 21411-21414, see entire document, especially abstract.	7-29
Y	US, A, 4,981,784 (Evans et al.), 01 January, 1991, see entire document.	23-27
Y	Cell, Volume 67, issued 15 November 1991, I.Rebay et al., "Specific EGF repeats of Notch mediate interactions with Delta and Serrate: Implications for Notch as a Multifunctional Receptor", Pages 687-699, see entire document.	10-15,19-29
Y	Development, Volume 111, issued 1991, U. Thomas et al., "The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs", pages 749-761, see especially pages 750-752.	1-29

* Special categories of cited documents:[16]

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search[2]	Date of Mailing of this International Search Report[2]
22 May 1992	17 JUN 1992
International Searching Authority[1]	**Signature of Authorized Officer[20]**
ISA/US	LORRAINE M. SPECTOR, PH.D.

Form PCT/ISA/210 (second sheet)(May 1986) &

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y	Genes and Development, Volume 4, issued 1990, R.J. Fleming et al., "The gene Serrate encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in Drosophila melanogaster", pages 2188-2201, see especially p. 2192.	1-29
Y	Annual Rev. Cell Biol. Volume 7, issued 1991, S. Artavanis-Tsakonas et al., "The Notch locus and the cell biology of neuroblast segregation", pages 427-452, see especially p. 446.	1-29

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE[1]

1. ☐ Claim numbers _, because they relate to subject matter (1) not required to be searched by this Authority, namely:

2. ☐ Claim numbers _, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out (1), specifically:

3. ☐ Claim numbers _, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING[2]

This International Searching Authority found multiple inventions in this international application as follows:

Please See Attached Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

1-29,56-63 (Telephone Practice)

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Search Authority did not invite payment of any additional fee.

Remark on protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]

FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS

VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claims 1-29 and 56-59, drawn to Serrate protein, derivatives, fragments and chimers thereof, as well as a method of making same. Classes 530, subclass 387 and 435, subclass 69.1
II. Claims 30-33, drawn to antibodies and fragments thereof. Class 424, subclass 85.8.
III. Claims 34-55, drawn to nucleic acids and transformed cells. Class 536, subclass 27.